

**УЗБЕКСКОЕ АГЕНТСТВО СВЯЗИ И ИНФОРМАТИЗАЦИИ**  
**ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ**  
**ТЕХНОЛОГИЙ**

**Факультет «РАДИОСВЯЗИ, РАДИОВЕЩАНИЯ И ТЕЛЕВИДЕНИЯ»**

**Кафедра «ТЕЛЕВИДЕНИЯ И РАДИОВЕЩАНИЯ»**

**АУДИОТЕХНИКА**  
**КОНСПЕКТ ЛЕКЦИЙ**

**Ташкент - 2011**

## 1-лекция. ЗВУК. ЗВУКОВЫЕ ЯВЛЕНИЯ.

План:

- Репрезентативная система
- Основные области применения электроакустики
- Звуковые волны
- Звуковое поле

Начнем издалека, поскольку наша задача не только выяснить, какие бывают микрофоны и динамики, но и почему они такие, для чего они и как их сделать лучше. Первоначально слуховая система использовалась, вероятно, как система безопасности. В отличие от зрения область чувствительности слуха составляет полный телесный угол. И сегодня система оповещения гражданской обороны основана именно на звуковом информировании: сирены и пр.

В дальнейшем на основе звуковых взаимодействий развилась речь - вторая сигнальная система. Это то, что сделало человека человеком.

**Репрезентативная система** - система восприятия через органы чувств и внутреннего моделирования окружающей действительности. В репрезентативной системе выделяются подсистемы, соответствующие основным органам чувств:

1. визуальная;
2. аудиальная;
3. кинестетическая (тактильные ощущения);
4. рациональная (дигитальная).

Прочие органы чувств несущественны. Подсистема, доминирующая при восприятии человеком окружающей действительности, называется **ключевой репрезентативной системой**. Подсистема, доминирующая при построении внутренних моделей действительности, называется **ведущей репрезентативной системой**.

На рис. 1.1 представлено распределение людей по группам в зависимости от ключевой репрезентативной системы [\[1\]](#). Аудиальная система здесь только на третьем месте, но это распределение характеризует в основном объем, а не качество, информации, поступающей по данному информационному каналу. Более адекватно важность информации отображает распределение по ведущим системам (рис. 1.2).

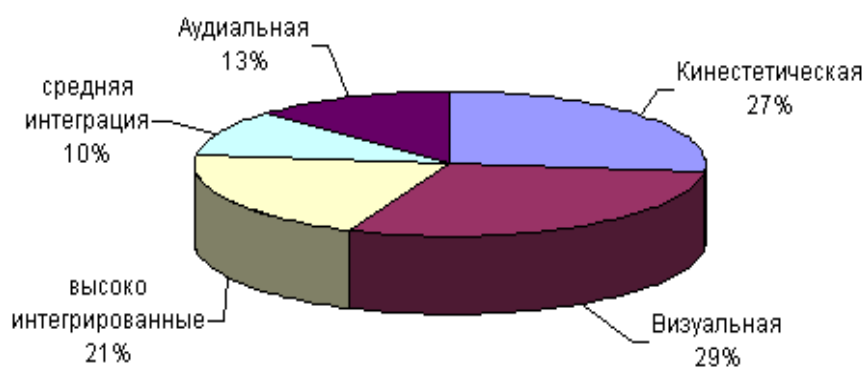


Рис. 1.1. Распределение людей по группам в зависимости от ключевой репрезентативной системы



Рис. 1.2. Распределение людей по группам в зависимости от ведущей репрезентативной системы

Можно видеть, что аудиальная система входит в состав ведущей репрезентативной системы у примерно 36-и процентов людей.

До того, как аудио канал стал использоваться в качестве второй сигнальной системы, он также использовался для передачи чувств и настроения посредством высоты, громкости и тембра звука. Это и стало, вероятно, объективной предпосылкой для возникновения музыки.

Никакое искусство ко времени появления звукозаписи так не нуждалось в технических средствах воплощения. Звук в отличие от масляных красок, бронзы и гранита вещь эфемерная и существует только в момент извлечения. В нотах может быть зафиксирована только мелодия, а часть музыки, связанная с непосредственной красотой звука и особенностями исполнительской интерпретации существует только в момент ее исполнения. Таким образом, запись можно рассматривать не только как "музыкальные консервы", но и как окончательно оформленное музыкальное произведение.

Первая электроакустическая система это, конечно, телефон. Изобретен Александром Беллом в 1876 г. В его честь названа универсальная единица измерения в логарифмических шкалах (и, в частности, силы звука) - Белл. В 1877 г. появилась первая система звукозаписи - фонограф - и с этого момента началась новая жизнь музыкального искусства. Фонограф считается изобретенным Томасом Эдисоном. Он не был электроакустической системой, но быстро выяснилось, что чисто механическая запись очень неудобная и неточная. Надо было очень близко садиться музыкантам к раструбу рекордера, с

молоточков пианино снимали смягчающие удар подушечки, в студии создавалась неестественно большая *реверберация* (стены обивали железом). И вот в 1898 году Вальдемар Паульсен (Дания) придумал переводить звуковые колебания в электрический ток, намагничивать стальную проволоку. Также в начале XX века происходил переход от чисто механической грамзаписи к записи с промежуточным преобразованием сигнала в электрический ток: механические колебания воздуха переводились в электрический сигнал, который затем усиливался вакуумными лампами и управлял электромагнитом, смещающим резец рекордера.

Изобретения телефона и устройств записи звука привели к возникновению новой отрасли науки и техники - электроакустики. **Электроакустика** - изучает технические средства перевода звуковых колебаний в электрический сигнал и обратного перевода электрического сигнала в максимально похожие звуковые колебания. Основные области применения электроакустики это:

- 1) звукоусиление;
  - 2) передача звука на большие расстояния (радиовещание, составляющая телевизионного вещания);
  - 3) запись звука с целью хранения и последующего воспроизведения.
- Звукозаписью** чаще называют процесс, а результат звукозаписи называется **фонограммой**.

Исходя из определения электроакустики, курс делится на два больших раздела.

1. Изучение объекта воспроизведения (первоначальных звуков) и условий его похожего воспроизведения (специфика распространения звука и особенностей слуха человека).

2. Методы перевода звукового поля в электрическую форму и возбуждения похожего звукового поля в другом помещении.

Дадим несколько определений.

**Звук.** Слово "звук" определяет два понятия: первое - звук как физическое явление; второе - звук как ощущение.

1. В результате вибрации (колебания) какого-либо упругого тела, например струны, возникает волнообразное распространение колебаний воздушной среды. Источником звука является колеблющееся тело. Оно приводит в колебательное движение прилегающие к нему частицы упругой среды (как правило, воздуха), которые заставляют колебаться соседние частицы и т.д. Процесс распространения колебаний частиц упругой среды называют **звуковой волной**.

2. Звуковые волны улавливаются слуховым органом и вызывают в нем раздражение, которое передается по нервным волокнам в головной мозг, возбуждая ощущение звука.

**Звуковое поле** - одна из форм существования материи, проявляется в виде кинетической энергии колеблющихся материальных тел, а также звуковых волн в твердой, жидкой и газообразной средах, обладающих упругой структурой.

**Первичное помещение** - помещение, в котором источниками звука создается звуковое поле (называемое первичным), подлежащее копированию при воспроизведении электроакустическими устройствами.

**Помещение прослушивания** - помещение, в котором прослушивается создаваемое электроакустическими преобразователями звуковое поле.

Вопросы:

- В репрезентативной системе какие выделяются подсистемы?
- Что изучает электроакустика?
- Основные области применения электроакустики?
- Понятие слово звук?
- Что называют звуковой волной?

## 2-Лекция. Параметры звукового сигнала.

План:

- Интенсивность звука
- Величины слухового ощущения
- Абсолютный порог слышимости
- Цифровой звук

Звук – это упругое колебание среды, он распространяется в среде с помощью волн давления посредством колебания атомов и молекул. Как и любая волна, звук характеризуется скоростью, амплитудой и частотой [5]. Воздействуя на слух, звук вызывает раздражение, которое создает у человека субъективный эффект – ощущение.

Интенсивность звука  $I$  определяется как среднее количество звуковой энергии, проходящей в единицу времени через единицу поверхности:

$$I = \frac{P^2}{Q_0}, \text{ Вт/м}^2 \quad (2.1)$$

где  $Q_0$  – удельное сопротивление воздуха  $Q_0 = 1,23 \text{ кг/м}^3$

При исследовании восприятия изменения интенсивности звука было установлено, что одинаковые относительные изменения интенсивности вызывают одинаковые абсолютные изменения слухового ощущения, то есть слуховое ощущение  $E$  пропорционально логарифму раздражающей силы:

$$E = a \lg I + C. \quad (2.2)$$

На пороге слышимости  $I = I_{\text{пс}}$  слуховое ощущение равно нулю, поэтому

$$C = -a \lg I_{\text{пс}} \quad (2.3)$$

откуда

$$E = a \lg(I/I_{\text{пс}}) \quad (2.4)$$

Для оценки величины слухового ощущения была предложена единица под название "бел" ( $a = 1$ ). Эта единица соответствует десятикратному отношению интенсивностей, поэтому была введена более мелкая единица – "децибел" ( $a = 10$ ).

$$E = 10 \lg(I/I_{\text{пс}}), \text{дБ} \quad (2.5)$$

Вследствие логарифмического закона восприятия звука и широкого диапазона слышимости звуков было введено понятие уровня интенсивности

$$N(I) = 10 \lg(I/I_0), \text{дБ} \quad (2.6)$$

За нулевой уровень условились принимать интенсивность  $I$  близкую к порогу слышимости для нормального слуха на частоте 1 кГц ( $I_0 = 10^{-12} \text{ Вт/м}^2$ ).

Нулевой уровень по звуковому давлению  $P$  равен  $2.04 \cdot 10^{-5} \text{ Па}$ . Интенсивность и звуковое давление связаны квадратичной зависимостью, поэтому уровень звукового давления  $N$  определяется равенством:

$$N = 20 \lg(P/P_0) \quad (2.7)$$

Порог слышимости – это минимальное звуковое давление, при котором еще существует слуховое ощущение. На рис. 2.1 показана зависимость порога слышимости от частоты, которая была выявлена опытным путем.

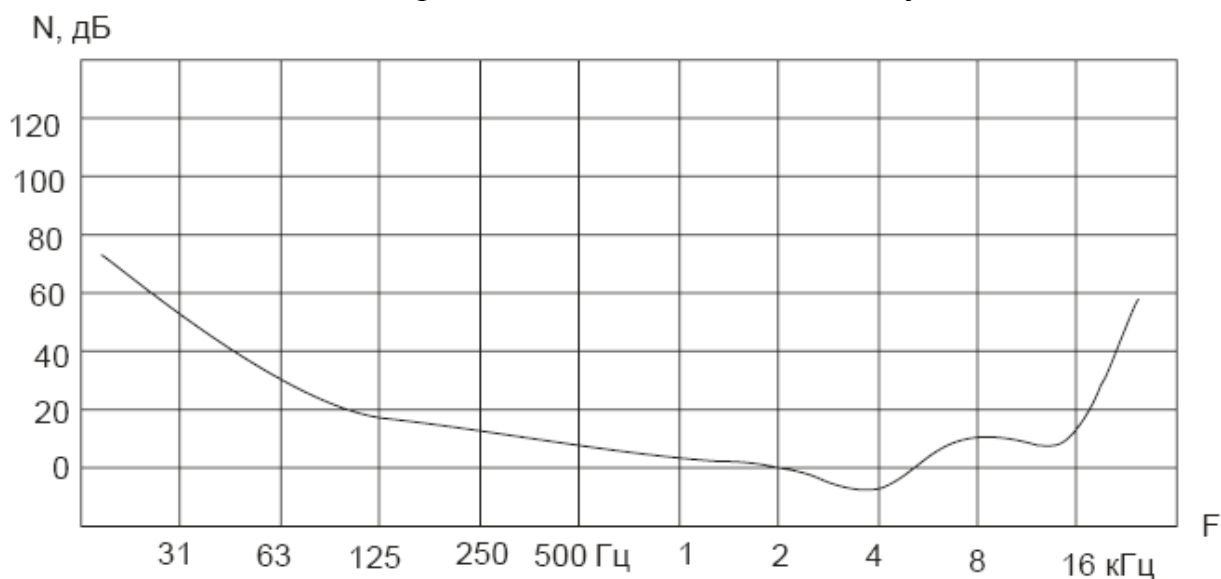


Рис. 2.1. Абсолютный порог слышимости

### Критические полосы слуха

При восприятии звука слуховой аппарат человека разделяет его на частотные группы, называемые критическими полосами. Эта особенность является одним из важнейших свойств слуха. Стандартами MPEG становлены нижние границы критических полос слуха: 20, 100, 300, 400, 510, 630, 770, 920,

1080, 1270, 1480, 1720, 2000, 2320, 2700, 3150, 3700, 4440, 5300, 6400, 7700, 9500, 12000, 16000 Гц.

В области до 500 Гц ширина частотных групп не зависит от средней частоты шума и равна примерно 100 Гц. В области выше 500 Гц она увеличивается пропорционально средней частоте (рис. 2.2).

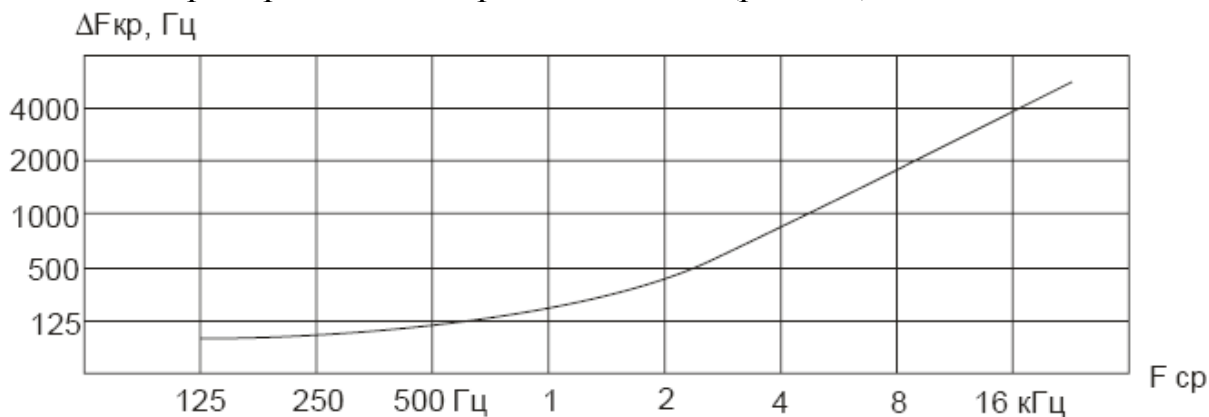


Рис. 2.2. Зависимость ширины критической полосы слуха от ее средней частоты

Критические полосы слуха оказалось удобно использовать в качестве единицы субъективной высоты тона, которую назвали барк. В настоящее время в психоакустических моделях слуха вместо частотных шкал используются 45 шкалы высоты тона в барках. На рис. 2.3 приведен график перевода частотной шкалы высоты тона в октавах в шкалу высоты тона в барках.

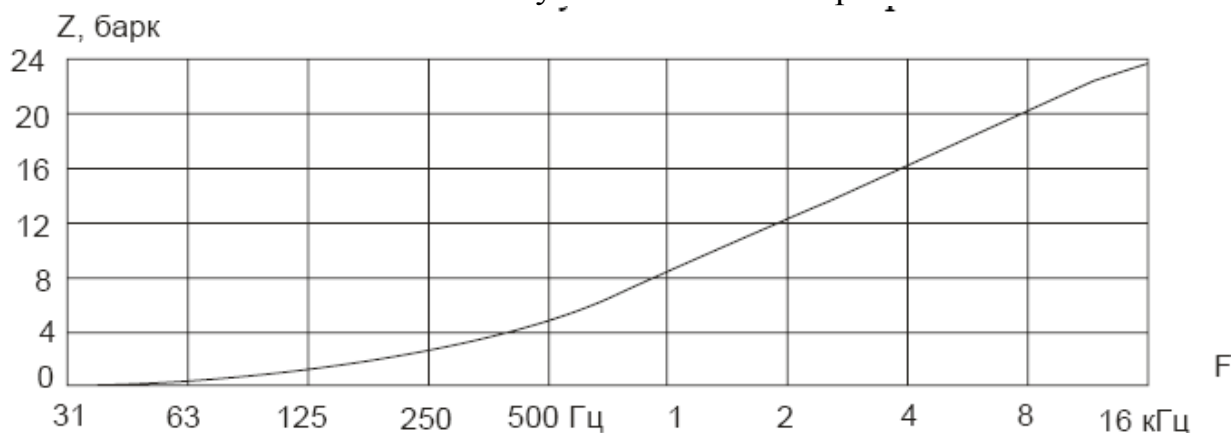


Рис. 2.3. График перехода от частотной шкалы высоты тона в октавах к шкале высоты тона в барках

### Пороги слышимости при маскировке

Повышение порога слышимости одного тона (или сигнала) при одновременном воздействии другого тона (шума или сигнала) называют маскировкой. Различают частотное и временное маскирование. На рис. 2.4 показаны кривые маскировки тона, когда мешающим звуком является тоже тон. По этим кривым определяют порог слышимости на конкретных частотах маскируемого тона, и рассчитывается коэффициент маскировки  $K_m$ , дБ

$$K_m = N_{ПС} - N_{АПС} \quad (2.8)$$

Маскировка максимальна, когда частоты  $T F$  и  $M F$  близки, величина маскировки увеличивается по мере возрастания уровня мешающего тона  $M N$ .

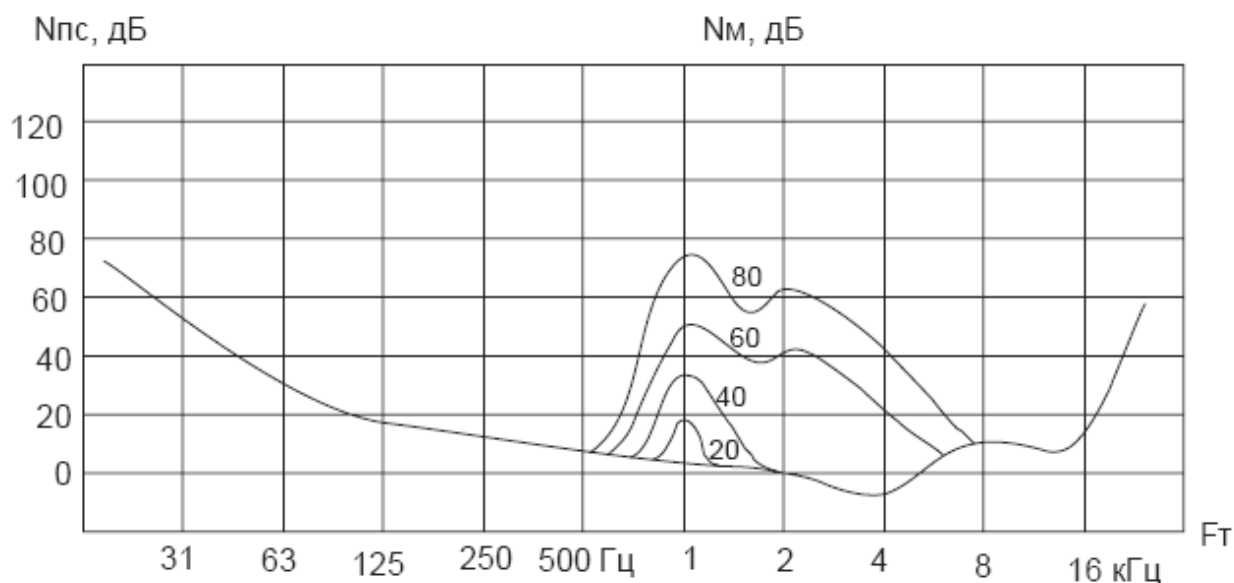


Рис. 2.4. Кривые порога слышимости тона с частотой  $T F$  при маскировке тоном с частотой  $M F = 1$  кГц и уровнем  $M N$ , дБ 46

На рис. 2.5 показаны кривые маскировки тона узкополосным шумом.

Они примерно такие же, как при маскировке тоном с некоторыми отличиями: нет биений и провала коэффициента маскировки. Узкополосный шум маскирует не только тоны, но и широкополосные шумы квантования.

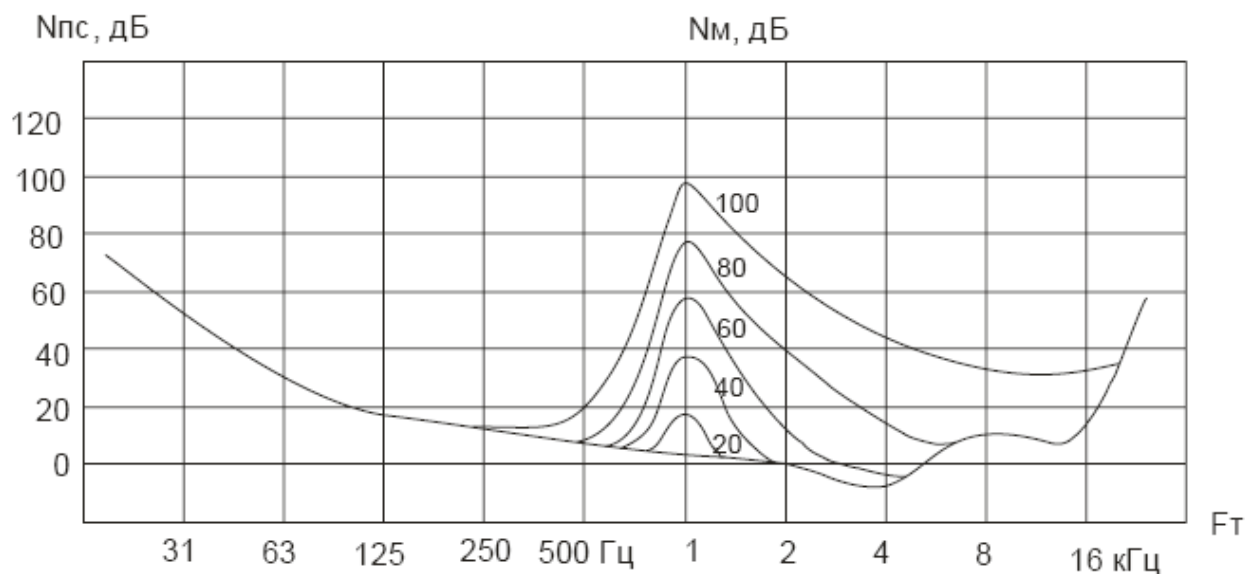


Рис. 2.5. Кривые порога слышимости тона с частотой  $T F$  при маскировке узкополосным шумом с центральной частотой  $M F = 1$  кГц и уровнем  $M N$ , дБ

На рис. 2.6 показан порог слышимости во временной области, при воздействии сигнала частотой  $f$  с уровнем интенсивности 60 дБ.



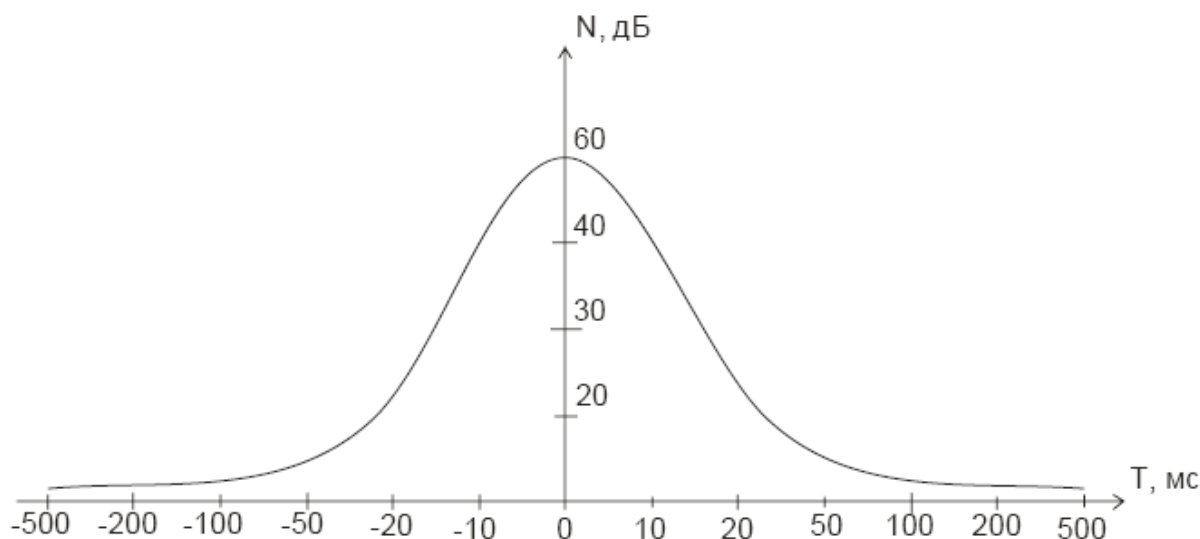


Рис. 2.6. Порог слышимости при временном маскировании

### Цифровой звук

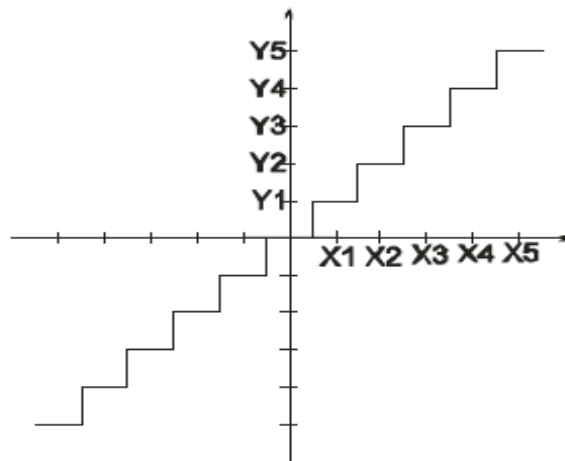
С записью звука в цифровом виде связано два действия: дискретизация и квантование. Дискретизация звука заключается в периодическом измерении аналогового сигнала и использовании полученных моментальных значений вместо исходной волны. Под квантованием понимается процесс получения предельно точных моментальных значений аналогового сигнала и последующего их округления.

**Дискретизация.** В соответствии с теоремой отсчетов В. А. Котельникова неискаженная передача непрерывного (аналогового) сигнала с полосой частот  $0..F_{\max}$  дискретной последовательностью его отсчетов возможна, если частота дискретизации  $Df$  связана с максимальной частотой  $\max F$  исходного сигнала соотношением:

$$\max f \geq 2F_{\max} Df \quad (2.9)$$

Если требуется передать синусоидальный сигнал с частотой 20 кГц, то требуемая частота его дискретизации должна быть более 40 кГц. Все сигналы, частота которых будет больше половины частоты дискретизации ( $2 Df > f$ ), при восстановлении будут интерпретированы как сигналы более низкой частоты.

**Квантование.** Квантование – это преобразование аналогового сигнала в ступенчатый сигнал с двоичным отсчетом уровней в квантах. При этой операции производится округление входного сигнала к принятой двоичной шкале квантователя. Процедуру квантования можно рассматривать как прохождение входного сигнала через устройство с амплитудной характеристикой ступенчатой формы (рис. 2.7), которая называется характеристикой (или шкалой) квантования. Если в пределах всей шкалы шаг квантования постоянен ( $\Delta x_i = \Delta x_{i-1}$  и  $\Delta y_i = \Delta y_{i-1}$ ), то квантование называют равномерным. Такой вид квантования удобен для начального представления звукового сигнала с целью последующей обработки<sup>1</sup>.



При использовании равномерного квантования получается сигнал в импульсно-кодовой модуляции (ИКМ)

Вопросы:

- Чем определяется интенсивность звука?
- Что называют порогом слышимости?
- Пороги слышимости при маскировке?
- Понятие цифрового звука?
- Что называется дискретизацией?
- Что называется квантованием?

### 3-Лекция. Основные виды алгоритмов сжатия аудиосигналов

К настоящему времени разработано множество алгоритмов кодирования общего назначения, которые выискивают повторяющиеся последовательности в двоичных данных. Если им удастся обнаружить такие последовательности, то они могут сжать данные, заменив последовательности кодами. Модуль, производящий декомпрессию (которому доступна информация о таких последовательностях) может произвести обратную операцию. Эти алгоритмы относятся к алгоритмам сжатия без потерь, так как восстановленные данные будут с точностью до бита идентичны исходным. Они могут быть использованы для компрессии любых данных, но при сжатии аудио сигналов оказываются мало эффективными [7].

Например, в методе Лемпеля-Зива-Уэлча (Lempel-Ziv-Welch, LZW), дефляции и методе Берроуза-Уиллера (Burroughs-Wheeler) алгоритм заключается в поиске длинных последовательностей байтов, встречающихся в файле несколько раз. Однако, в звуковых файлах длинных повторяющихся последовательностей обычно не много: в основном это обуславливается наличием шумов.

В алгоритме Хаффмана (Huffman) и арифметическом кодировании ищутся байты с определенными значениями (или пары байтов), попадающие в файл

чаще остальных. Как только удастся выделить такое значение, строится код, который тем короче, чем чаще встречается значение. Действительно, у звуковых файлов неравномерно распределены возможные значения байтов, поэтому алгоритмы данного типа могут достаточно хорошо обрабатывать звуковые файлы. Но так как звуковые файлы могут обладать очень большим объемом, эти методы также становятся неэффективными. Как правило, методы сжатия без потерь используются в современных стандартах сжатия на вторичном, дополнительном этапе сжатия с целью повышения коэффициента сжатия.

Еще одним недостатком алгоритмов сжатия без потерь является неоднородность компрессии, что не позволяет наложить ограничения на скорость передачи сжатых аудио потоков. Учет основных требований к качеству речевого сегмента и временной задержке, как правило, требует компромисса [5]. С увеличением сложности алгоритма для уменьшения требуемой полосы пропускания увеличивается время обработки в устройствах кодирования/декодирования. Хотя качество сигнала в основном напрямую связано со скоростью цифрового потока, сложные алгоритмы кодирования способны достичь более высоких отношений качества и скорости. Сжатие речевого сегмента (РС) может быть как без потерь (архивация), так и с потерями. Причем в последнем случае это кодирование можно подразделить на три вида [2,3]:

1. Кодирование непосредственно реализации РС (Wave Form Codec);
2. Измерение, кодирование и передача на приемную сторону параметров РС, по которым уже на приемной стороне производится синтез этого (искусственного) РС. Такие системы называют вокодерными (Source Codec);
3. Гибридные способы кодирования, т.е. сочетание первого и второго способов кодирования. Под кодированием подразумевается преобразование РС в некоторый «другой» сигнал, который можно представить с меньшим числом разрядов, что в итоге повысит скорость передачи данных. Одним из видов такого кодирования является дифференциальная импульсно-кодовая модуляция (ДИКМ), о которой и пойдет речь в дальнейшем.

Кодеры формы волны аппроксимируют изменение сигнала во времени. Они требуют наибольших скоростей передачи, но имеют наилучшие показатели качества воспроизведенной речи. При параметрическом компандировании моделируется процесс речеобразования человека. В кодере из речевого сигнала вычисляются определенные параметры, передаваемые к декодеру, в котором они применяются для восстановления формы исходного сигнала. Использование исключительно параметрических методов приводит к потере натуральности звучания голоса и большой чувствительности к фоновым шумам. Вокодерные преобразования отличаются наименьшими требованиями к полосе пропускания. Один из способов повышения эффективности использования полосы пропускания состоит в применении *гибридных методов*, основанных на принципах линейного предсказания и объединяющих параметрическое компандирование и кодирование формы волны. Большинство гибридных кодеров используют замкнутое кодирование (метод «анализ через

синтез») на передающей стороне, что позволяет подкорректировать определенные параметры посредством сравнения результата синтеза с оригиналом. Это, безусловно, увеличивает время обработки, но обеспечивает лучшие показатели при передаче.

На рис.2.8 - 2.10 отображены основные операции для различных методов кодирования.

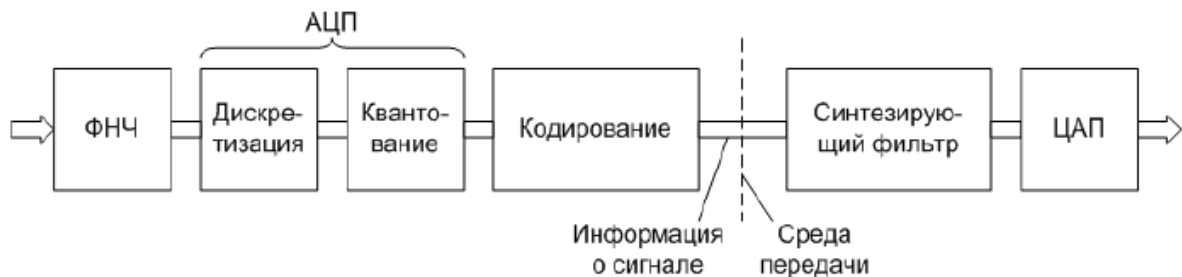


Рис.2.8. Основные операции кодирования формы волны

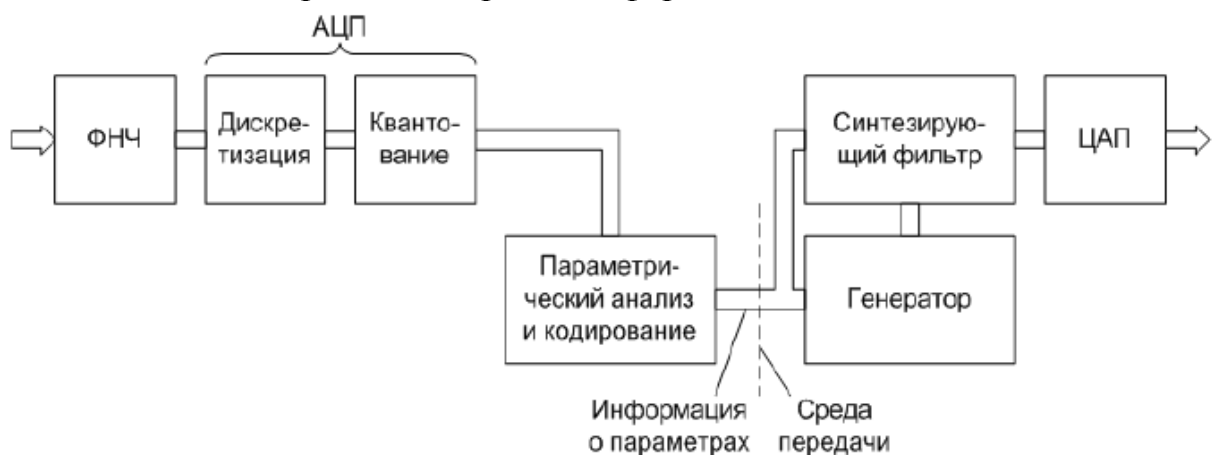


Рис.2.9. Основные операции параметрического компандирования

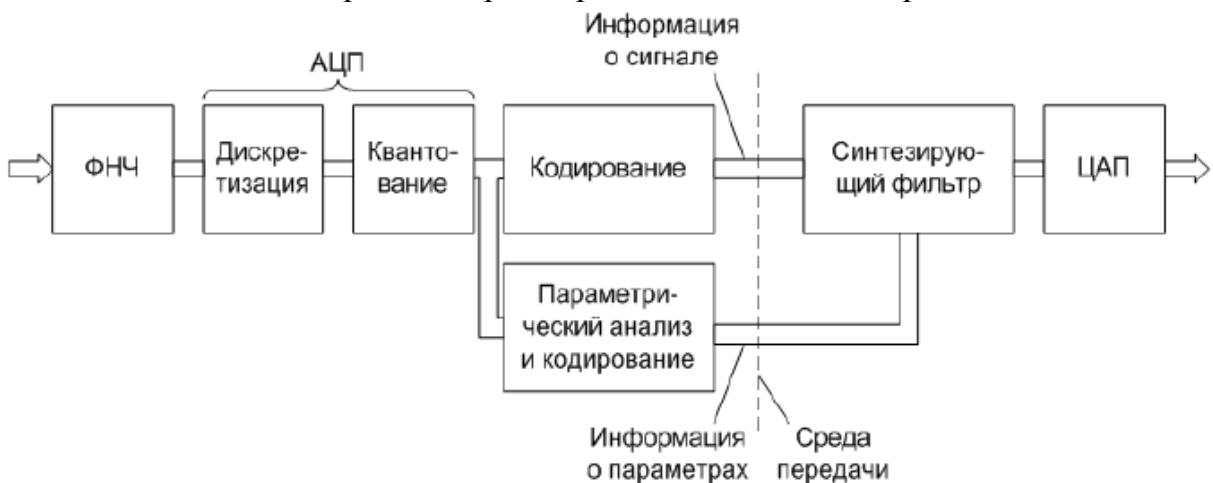


Рис.2.10. Основные операции кодирования гибридными методами

В табл.2.1 представлены наиболее распространенные алгоритмы и области их применения [5]. Алгоритмы указаны в порядке убывания битовой скорости потока. Как правило, определяющими для выбора метода кодирования являются такие показатели, как:

- качество голоса по пятибалльной шкале экспертных оценок MOS (Mean Opinion Score, Рекомендация ITU-T P.800);
- задержка алгоритма;

- помехоустойчивость;
- степень ухудшения качества сигнала при квантовании QDU (Quantization Distortion Units);
- распространенность, поддержка производителями оборудования и др.

Таблица 2.1

Алгоритм	Скорость, кбит/с	Стандарт	Год	Приложение
Аппроксимация формы речевой волны				
PCM <sup>2</sup>	64, 56, 48	ITU-T G.711	1960	Общественные телефоны
SB-ADPCM <sup>3</sup>	64, 56, 48	ITU-T G.722	1986	Передача широкополосных сигналов
ADPCM <sup>4</sup>	32	ITU-T G.721	1984	Общественные телефоны

Таблица 2.1

Алгоритм Скорость, кбит/с

Стандарт Год Приложение

Аппроксимация формы речевой волны

PCM<sup>2</sup> 64, 56, 48 ITU-T G.711 1960 Общие телефоны

SB-ADPCM<sup>3</sup> 64, 56, 48 ITU-T G.722 1986 Передача широкополосных сигналов

ADPCM<sup>4</sup> 32 ITU-T G.721 1984 Общие телефоны

2 PCM (Pulse Code Modulation) – импульсно-кодовая модуляция.

3 SB-ADPCM (Sub-Band Adaptive Differential Pulse Code Modulation) – адаптивная

дифференциальная импульсно-кодовая модуляция с делением на поддиапазоны.

4 ADPCM (Adaptive Differential Pulse Code Modulation) – адаптивная дифференциальная импульсно-кодовая модуляция.

ADPCM	40, 32, 24, 16	ITU-T G.726	1984	Общественные и цифровые беспроводные телефоны
Гибридные методы кодирования				
LD-CELP <sup>5</sup>	16	ITU-T G.728	1992	Общественные телефоны, видеотелефоны
RPE-LTP-LPC <sup>6</sup>	13	ETSI GSM 06.10	1992	Европейские цифровые сотовые системы
CS-ACELP <sup>7</sup>	11.8, 8, 6.4	ITU-T G.729, G.729 Annex	1997	Передача речи в сетях Frame Relay, ATM, в системах телесвязи Франции
MP-MLQ <sup>8</sup>	6.3	ITU-T G.723.1	1996	Передача речи в видеотелефонии
VSELP <sup>9</sup>	5.6	ETSI GSM 06.20		Европейские цифровые сотовые системы
ACELP <sup>10</sup>	5.3	ITU-T G.723	1996	Передача речи в видеотелефонии
Вокодерные преобразования				
LPC-1011	2.4	ANSI		Специальные системы

ADPCM 40, 32, 24,

16 ITU-T G.726 1984 Общие и цифровые беспроводные телефоны

Гибридные методы кодирования

LD-CELP5 16 ITU-T G.728 1992 Общественные телефоны,  
видеотелефоны

RPE-LTP/LPC6

13 ETSI GSM 06.10 1992 Европейские цифровые сотовые  
системы

CS-ACELP7 11.8, 8,  
6.4

ITU-T G.729,  
G.729

Annex 1997

Передача речи в сетях Frame Relay, ATM, в системах телесвязи Франции  
MP-MLQ8 6.3 ITU-T G.723.1 1996 Передача речи в видеотелефонии VSELP9 5.6  
ETSI GSM 06.20 Европейские цифровые сотовые системы ACELP10 5.3 ITU-T  
G.723 1996 Передача речи в видеотелефонии Вокодерные преобразования LPC-  
1011 2.4 ANSI Специальные системы В табл.2.2 приведены данные по  
соответствию качества речи, MOS, задержек передачи и типов каналов,  
удовлетворяющих предъявленным требованиям [5].

Таблица 2.2.2

Таблица 2.2.2

Качество	Лучшее	Хорошее	Среднее	Плохое	Стандарт
MOS	> 4.5	4–4.5	3.5–4	3–3.5	ITU-T P.800, P.830
Задержка, мс	< 150	< 250	< 350	< 450	ETSI TS 101 329
	< 150	< 260	< 400	> 400	ITU-T G.114
Тип канала	ТфОП	Спутниковый	ТфОП + спутниковый	Допустимо для VoIP	

5 LD-CELP (Low Delay Code Excited Linear Prediction) – линейное  
предсказание с кодовым возбуждением и малой задержкой. 6 RPE-LTP-LPC  
(Regular Pulse Excitation Long Time Prediction Linear Predictive Coding) –  
кодирование на основе линейного предсказания с долговременным  
предсказанием с регулярным импульсным возбуждением. 7 CS-ACELP  
(Conjugate Structure Algebraic Code Excited Linear Prediction) – линейное  
предсказание сопряженной структуры с алгебраическим возбуждением. 8 MP-  
MLQ (Multi Pulse Maximum Likelihood Quantization) – метод квантования по  
максимуму правдоподобия. 9 VSELP (Vector Sum Excited Linear Prediction) –  
линейное предсказание с векторным возбуждением.

10 ACELP (Algebraic Code Excited Linear Prediction) – линейное  
предсказание с алгебраическим возбуждением. 11 LPC (Linear Predictive  
Coding) – кодирование на основе линейного предсказания.

На рис.2.11 изображены сглаженные зависимости оценок MOS от  
требований к битовой скорости потока, построенные по усредненным

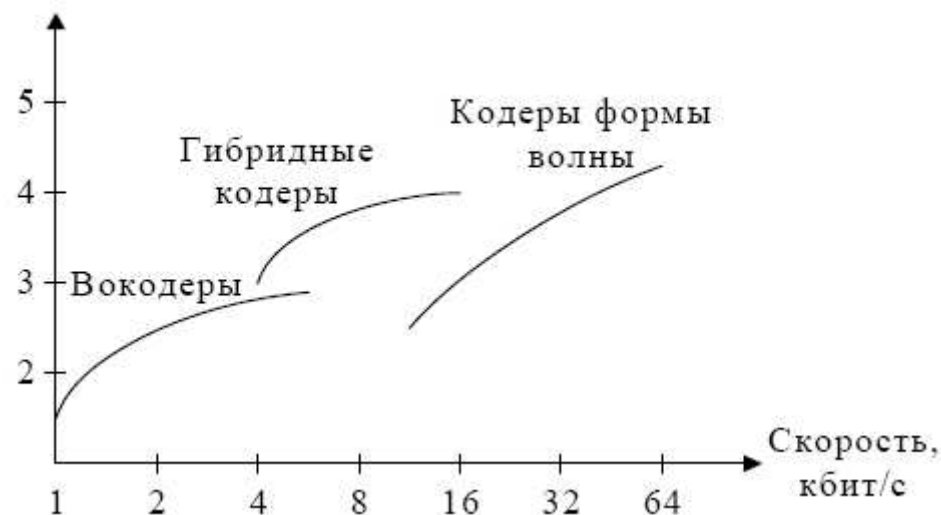


Рис.2.11. Зависимость оценок MOS от скорости потока для кодеров формы волны, вокодеров и гибридных кодеров

Следует отметить, что значения MOS можно встретить во многих информационных источниках, при этом отклонения составляют не более 0.4 балла, что допустимо, поскольку «хорошая» связь или «плохая» – это субъективная оценка, зависящая от ожиданий абонентов, их капиталовложений и других факторов.

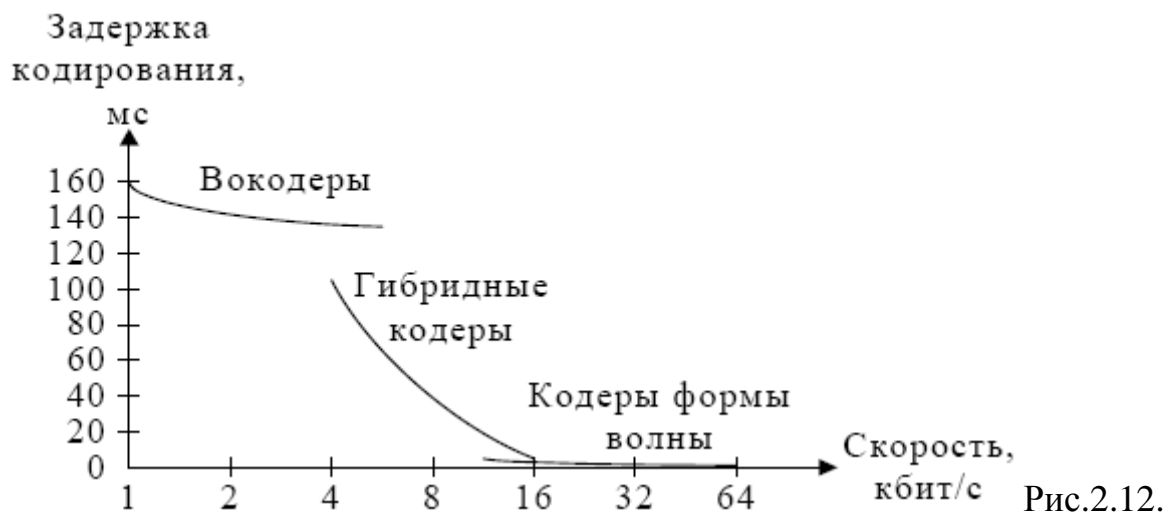
Указанные в табл.1.2 значения задержек следует считать ориентировочными, можно встретить и другие.

На количественный показатель задержки оказывают воздействие [2]:

- алгоритмы кодирования/декодирования информации;
- сеть;
- операционная система;
- буфер устранения джиттера.

Следует отметить, что только среда передачи в среднем задерживает сигнал на 10–150 мс в зависимости от длины и типа каналов связи. Как правило, более сложные алгоритмы кодирования обеспечивают лучшее сжатие при практически неизменном качестве речи. Становится очевидным, что для уменьшения битовой скорости, а, следовательно, и составляющих сетевой задержки, неизбежно увеличение задержки алгоритмов кодирования.

Безусловно, временная задержка кодирования зависит от быстродействия устройства, выполняющего преобразование. Поэтому представленные ниже графики, по мнению автора работы, правильнее рассматривать с точки зрения относительных значений, а не абсолютных. На рис.2.12 изображены сглаженные зависимости общих задержек алгоритмов от битовой скорости потока.



Зависимость задержки кодирования от скорости потока для кодеров формы волны, вокодеров и гибридных кодеров. Задержки декодирования могут существенно изменяться в зависимости от организации буфера устранения джиттера. В области телекоммуникаций джиттером называется первая производная задержки прохождения данных по времени.

#### 4-Лекция. Форматы сжатия звуковых сигналов





Рис. 1.2 Структурная схема кодера MPEG стандарта ISO/IEC 11172-3, Layer-1 и Layer-2

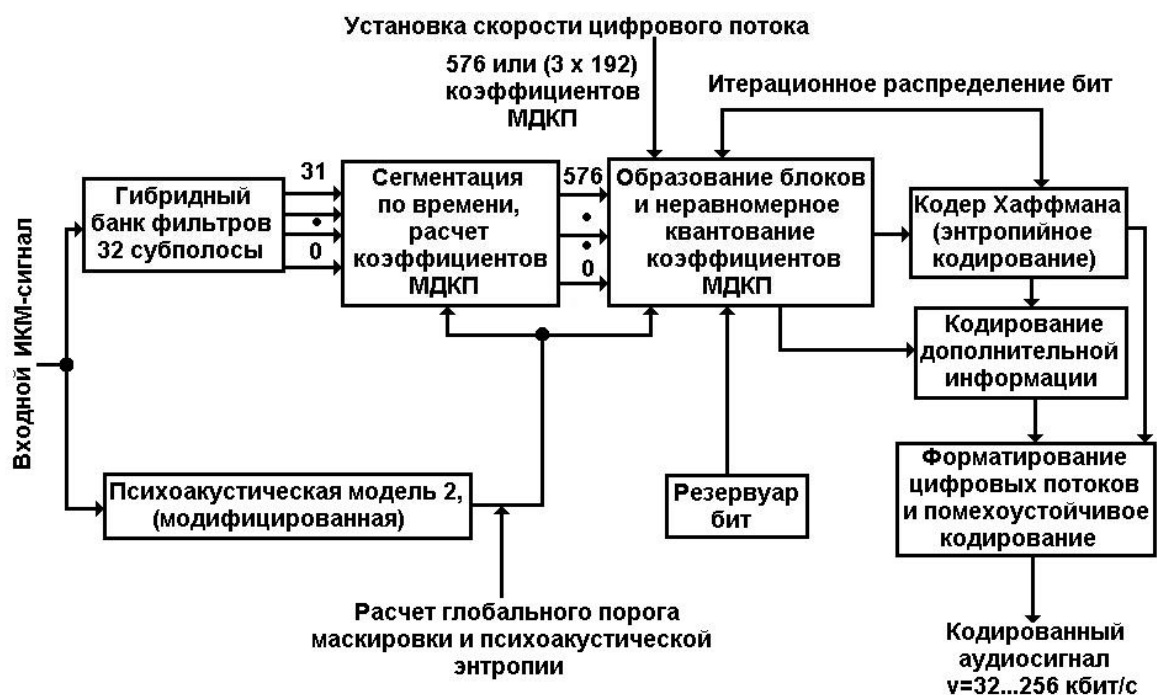


Рис. 1.3 Структурная схема кодера MPEG стандарта ISO/IEC 11172—3 Layer-3

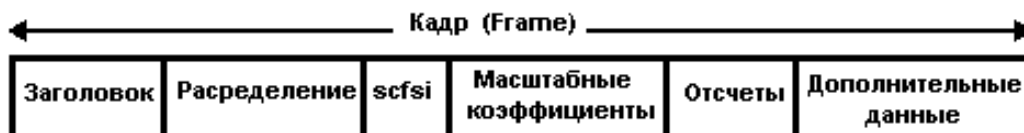


Рис. 1.4 Формат пакета MPEG-1

Каждый аудиокادر состоит из четырех независимых частей:

**1. Заголовок (Header).** Он представляет собой информацию, содержащуюся в первых 32 битах (4 байта) аудиокадра и обеспечивает синхронизацию и перенос остальной информации, требующейся для декодирования потока.

**2. Error\_check** (проверка ошибок) — это опциональная часть потока, которая содержит циклическую проверку на избыточность (Cyclical Redundancy Check, crc\_check), она представляет из себя 16-битное слово для проверки на четность и обеспечивает обнаружение ошибок в кодированном потоке.

**3. Audio\_data** (аудиоданные) — Звуковая информация, которая содержит кодированные отсчеты и информацию, необходимую для их декодирования.

**4. Ancillary\_data** (дополнительные данные) содержат биты, которые могут быть добавлены пользователем. Количество используемых дополнительных бит (no\_of\_ancillary\_bits) должно быть вычтено из общего количества бит в кадре. Так как длина кадра постоянна, то это вычитание возможно для кодирования отсчетов аудио и может иметь сильное влияние на качество аудио.

Для того чтобы грамотно выбрать речевой кодек, достаточно представления об используемом в нем методе (на котором базируется алгоритм кодирования) и о процессе согласования сигнала, полученного после цифровой обработки, с цифровым каналом связи.

Поскольку рассматриваемые методы кодирования являются методами сжатия звука с потерями, то при восстановлении (декодировании) звукового сигнала наблюдаются искажения сигнала. В качестве примера подобных искажений на рис. 2.22. приведен фрагмент исходного звукового сигнала, а на рис. 2.23 и рис.2.245 – соответствующие ему фрагменты после восстановления кодеками MPEG AAC и GSM 6.10.

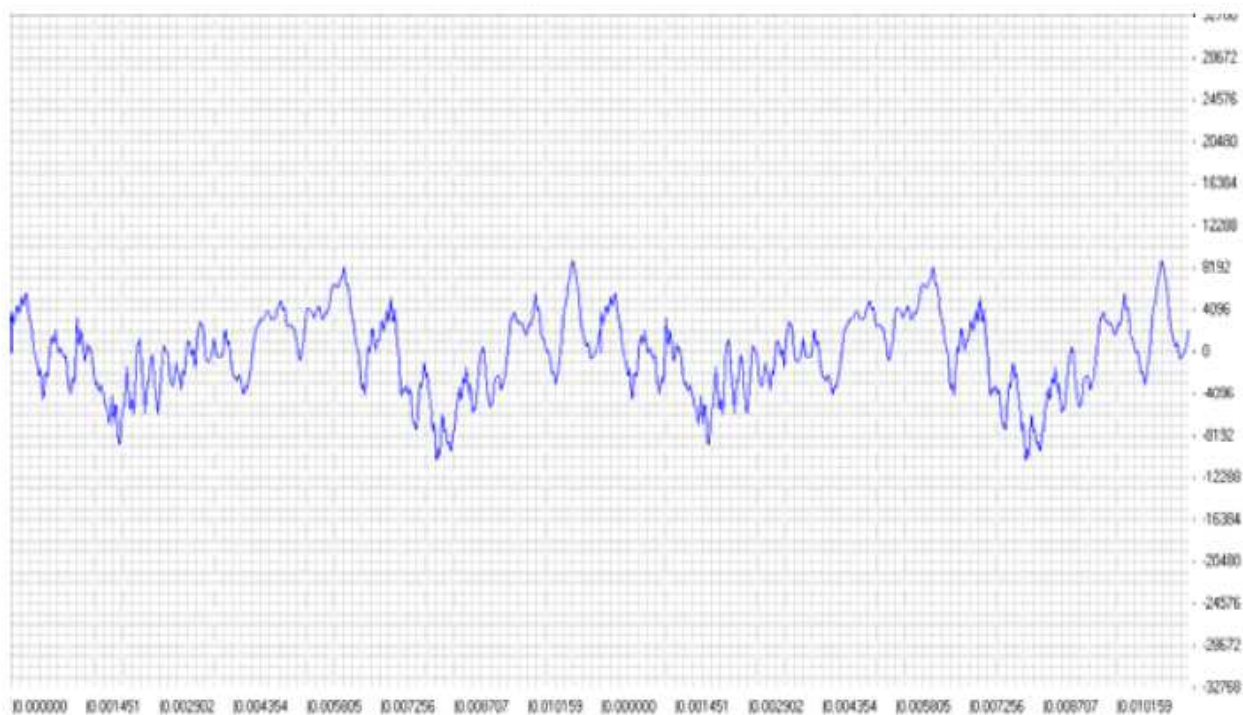


Рис.2.22. Фрагмент звукового сигнала  
(по оси X – время, мс, по оси Y – амплитуда)

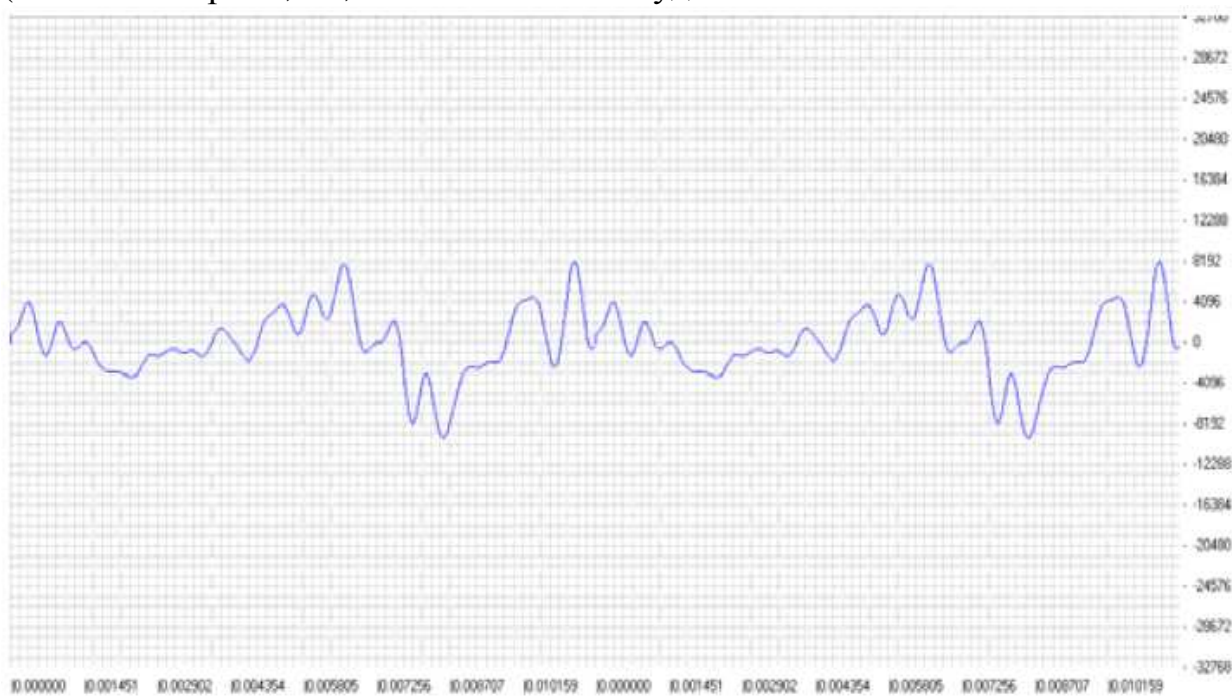


Рис.2.23. Фрагмент того же сигнала после декодирования кодеком MPEG AAC

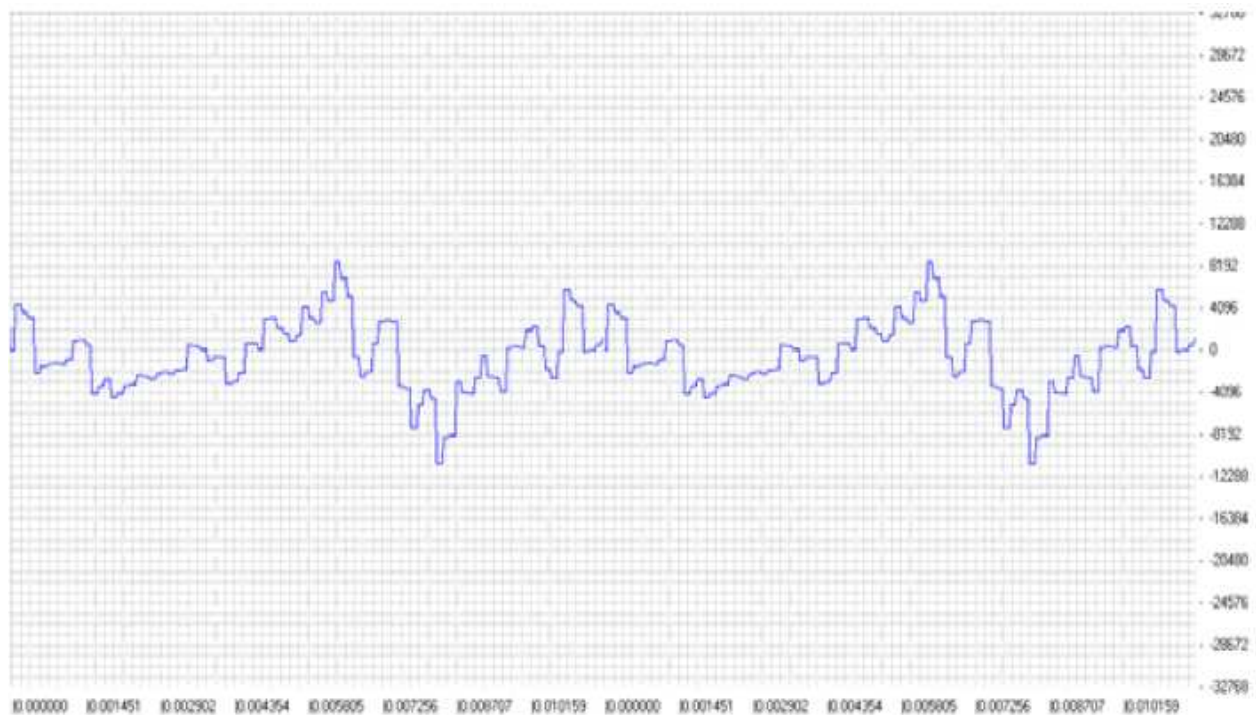


Рис.2.24. Фрагмент того же сигнала после декодирования кодеком GSM 6/10

Из рис.2.24 следует, что применение кодека MPEG AAC приводит к устранению высокочастотных компонент сигнала (Эффект сглаживания, а кодек GSM 6.10 вносит характерную ступенчатую структуру формы сигнала что связано с существенно меньшим размером базового фрейма при сжатии и отсутствием перекрытий между блоками.

## 5-Лекция. Качественные показатели методов кодирования речи.

Зависимость качества звучания речи (в баллах по шкале MOS) от скорости передачи при использовании различных методов кодирования показана на рис.2.25. По эффективности (т.е. соотношению битрейта к заданному качеству) методы кодирования можно разделить на три группы. К первой группе - низкоэффективных кодеров - относятся ИКМ и АДМ (адаптивная дельта-модуляция), которые обеспечивают удовлетворительное качество звука только при скоростях передачи выше 24 кбит/с. Кодеры второй группы - ОПА, МПК, ЛПКВ, АДИКМ - позволяют реализовать удовлетворительное и отличное качество звучания при скоростях 8-32 кбит/с. В третью группу входят ЛПКВ- и ЛПМВ-кодеры, обеспечивающие отличное качество при низких скоростях. В большинстве алгоритмов эффективного кодирования речи вероятность однократной ошибки составляет  $10^{-4}$ - $10^{-3}$  на символ. По зависимости помехоустойчивости от скорости передачи (рис. 2.26) кодеры можно разделить на две группы: в первую (ИКМ, АДМ, АДИКМ) входят алгоритмы с сильной зависимостью, во вторую (МПК, ЛПКВ, ОПА, ЛПМВ) - со слабой.

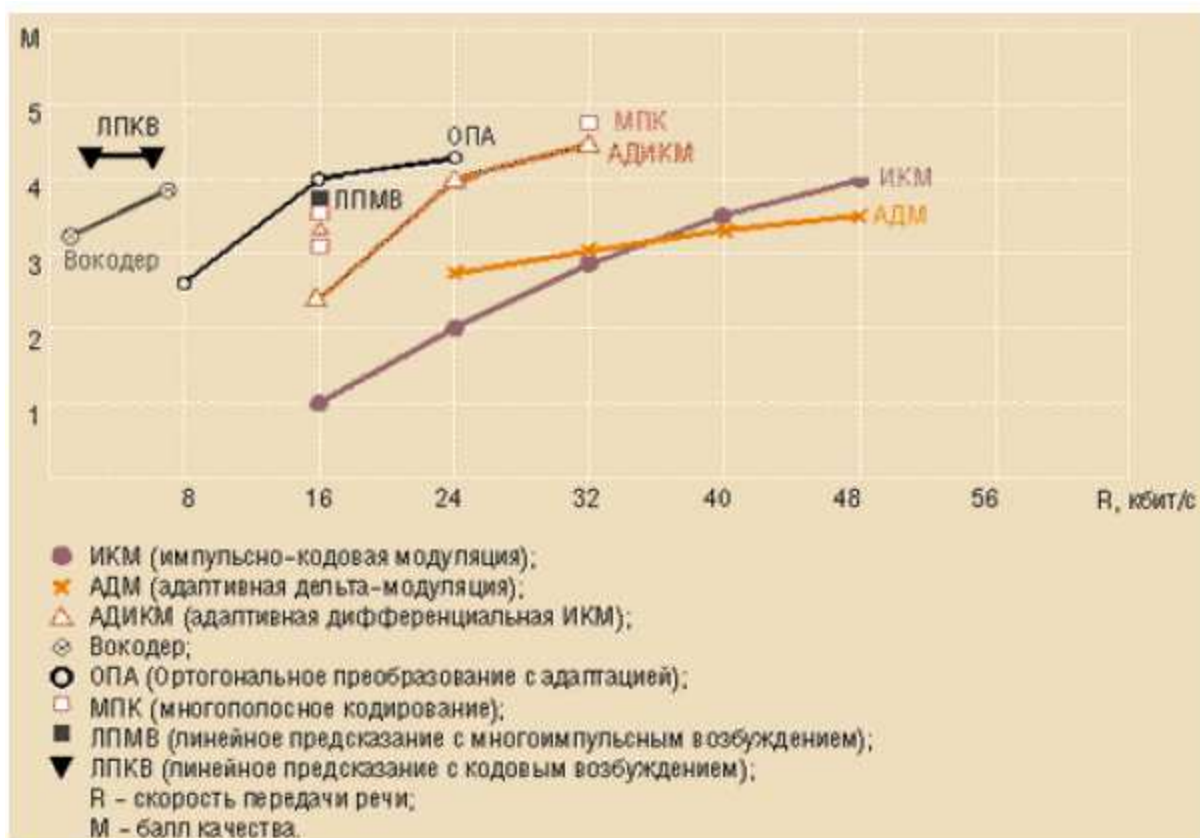


Рис. 2.25. Зависимость качества восстановления от битрейта R для различных методов кодирования

Под помехоустойчивостью подразумевается максимальная вероятность ошибки, при которой качество звучания сигнала (отношение сигнал/шум) снижается не более чем на 10% (в испытаниях по методу парных сравнений различия между сигналами составляют не более 20%).

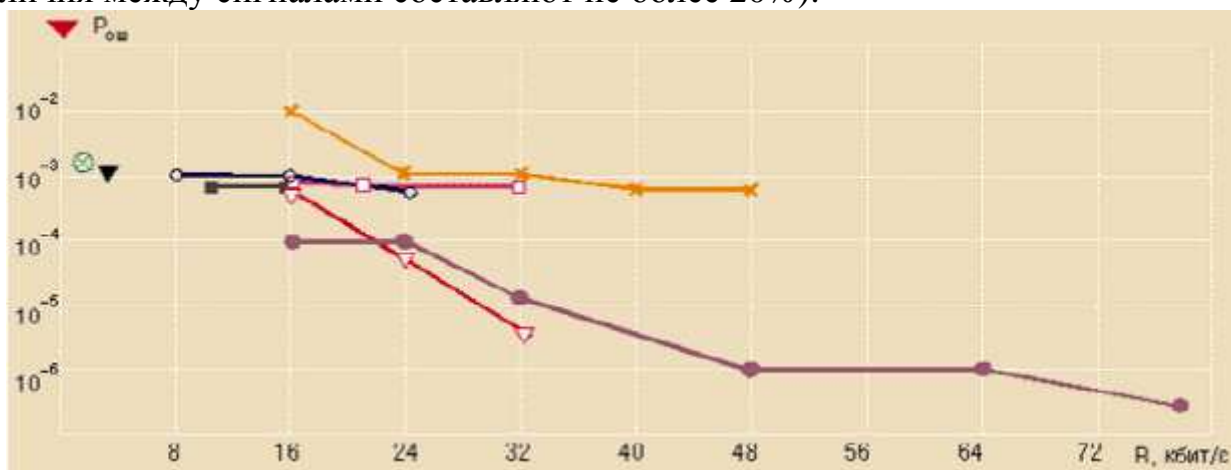


Рис. 2.26. Зависимость вероятности появления одиночной ошибки при декодировании (помехоустойчивости) кодеров от битрейта R (условные обозначения те же, что на рис. 1.6)

Необходимо отметить, что большинство кодеров предназначены для высококачественных цифровых каналов, но, несмотря на это, допустимая и максимально удерживаемая вероятности ошибки являются стандартными характеристиками всех эффективных кодеров.

**Выводы.** Рассмотрение существующих кодеков позволяет сформулировать ряд основных тенденций:

1. Доминирующее положение при построении низкоскоростных кодеков речи метода кодирования на основе линейного предсказания.

2. Возрастание доли адаптивных процедур обработки сигналов в современных системах кодирования речи.

3. Однозначная связь качества синтезированной речи на низких скоростях кодирования со степенью адаптации соответствующих кодеков речевых сигналов.

4. Наиболее перспективными являются алгоритмы сжатия типа CELP и его вариаций, а также векторного кодирования HVXC, положенные в основу методов MPEG AAC, и более современных стандартов MPEG.

5. Использование более сложных алгоритмов энтропийного кодирования (CABAC – контекстно-адаптивного арифметического кодирования) Таким образом, анализ степени адаптации современного парка кодеков указывает на перспективность перехода систем кодирования речи к многопараметрической адаптации в условиях априорной и текущей неопределенности в описании моделей речевого сигнала и внешней среды функционирования кодека.

Наибольшая компрессия достигается в методах, которые учитывают особенности человеческого слуха, использующих разбиение на поддиапазоны и последующего проведения в них анализа. Некоторые примеры таких методов приведены в табл. 2.3. Кодеки MPEG являются наиболее распространенными и используют наиболее перспективный метод сжатия.



Таблица 1.3.

Наименование алгоритма компрессии аудиоданных	Метод компрессии	Скорость передачи, кбит/с на канал	Величина компрессии	Области применения
ASPEC	Кодирование с преобразованием	64-192	1:6	ISDN
ATRAC	Субполосное кодирование с преобразованием	256	1:5	MiniDisk
MUSCAM	Субполосное кодирование	128-256	1:4	DAB (Digital Audio Broadcasting)
MUSICAM	Субполосное кодирование	128-256	1:4	DAB
MPEG-1, Layer 1 и Layer 2	Субполосное кодирование (MUSICAM)	32-448 (Layer 1) 32-384 (Layer 2)	1:4 (Layer 1) 1:6 (Layer 2)	DAB (Layer 2, 128-256 кбит/с), DBS (Direct Broadcast Satellite, Layer 2, 224 кбит/с), DCC (Digital Compact Cassete, Layer 1, 384 кбит/с)
MPEG-1, Layer 3	Субполосное кодирование с преобразованием	32-320	1:9	Internet-вещание
MPEG-2	Субполосное кодирование/ субполосное кодирование с преобразованием	32-384	> 1:9	Многоканальное стереофоническое вещание
MPEG-2 AAC	Субполосное кодирование с преобразованием	16-384	1:15	Многоканальное стереофоническое вещание
MPEG-4	Субполосное кодирование с преобразованием/ параметрическое кодирование	2-64	-	Мультимедиа приложения
Dolby AC-3	Кодирование с преобразованием	32-384	1:13	Кинематограф, HDTV, спутниковое вещание
Гибридное	Субполосное кодирование с преобразованием/	32-64	1:(15-20)	Радиовещание, хранение информации

## 6-Лекция. Современные многоканальные системы

**Dolby Digital (AC-3, ATSC A/52)** (Долби Диджитал) — система пространственного звуковоспроизведения, разработанная фирмой «Dolby Laboratories, Inc.» («Dolby Labs»), руководимой Реєм Долби, пионером аудио- и видеопромышленности.

Формат стандартизирован *Advanced Television Systems Committee*, ему присвоен код A/52, *Dolby Digital (DD)* является торговой маркой.

### ***Dolby Digital Live***

Dolby Digital Live (DDL) - технология кодирования многоканального (5.1) аудиосигнала в формат AC3 в реальном времени, предложенная компанией Dolby Technologies. Предназначена для передачи многоканального звука из игр и иных приложений на ресивер по интерфейсу S/PDIF(оптическому или коаксиальному).

### ***Dolby Digital Plus***

Компании MIPS Technologies и Dolby Laboratories представили новую технологию построения звука для устройств, поддерживающих воспроизведение видео и аудио высокой четкости, например HD DVD и Blu-ray проигрыватели. Аудио-технология получила название Dolby Digital Plus и сможет использоваться в микропроцессорных ядрах MIPS32.

### ***Dolby TrueHD***

Dolby TrueHD является одним из первых двух форматов несжатого звука (сжатого без потерь), доступных только для оптических HD-плееров. Несмотря на то, что кодек Dolby TrueHD является опциональным, данный формат широко поддерживается плеерами и дисками Blu-ray (гораздо больше, чем DD+, который часто отсутствует на дисках Blu-ray). Dolby TrueHD использует алгоритм сжатия без потерь Meridian Lossless Packing (MLP). Цифровой поток Dolby TrueHD может вмещать до 14 отдельных звуковых каналов, но на практике работает с 6 (5.1) или 8 (7.1) каналами.



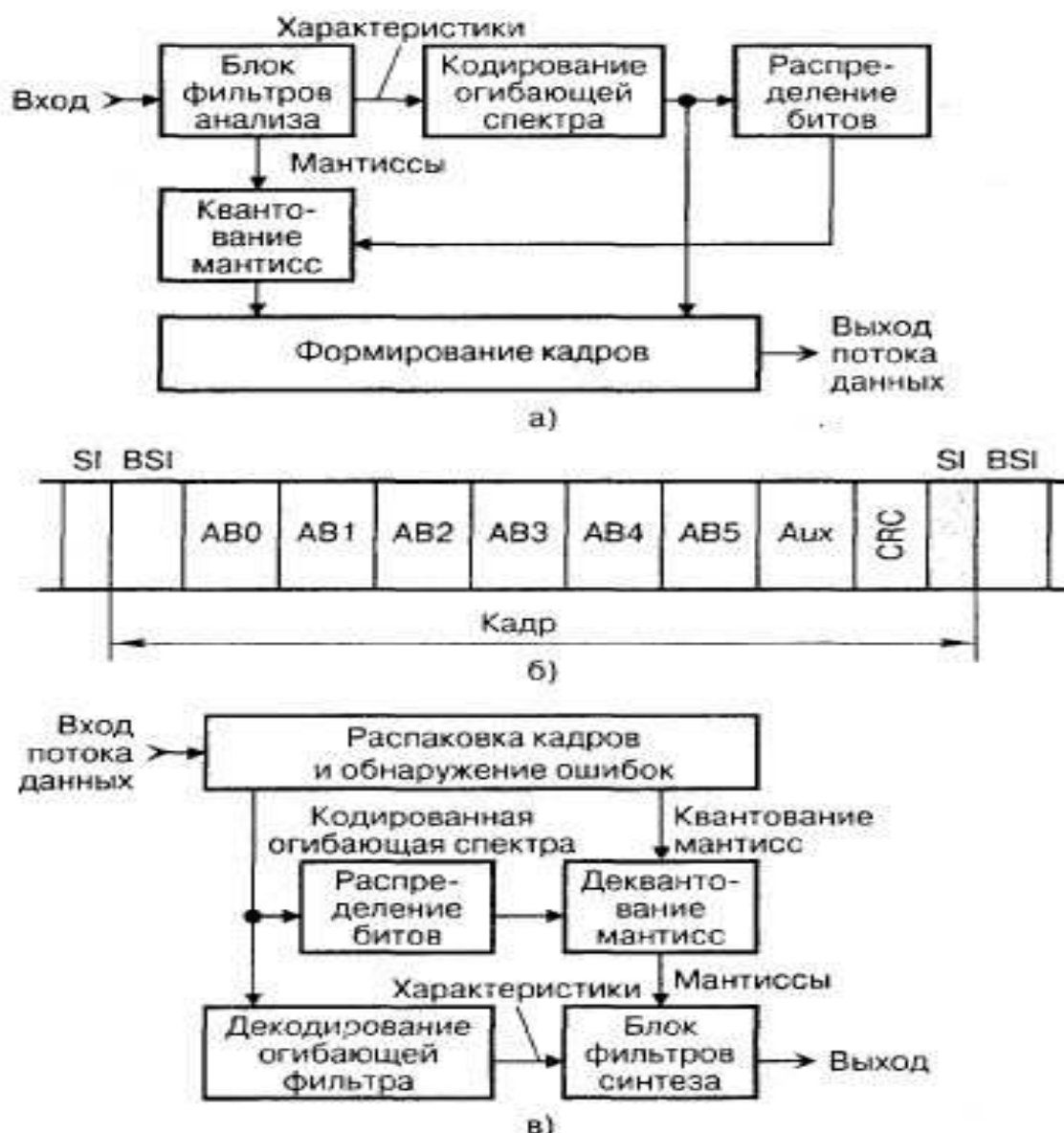


Рис. 1.4. Стандарт Dolby AC-3: кодер (а), структура кадра (б), декодер (в)

Передать пространственные звуковые ощущения - задача со многими вариантами решений. Мы не раз сталкивались с идеализированием звуковых систем, разговорами о полноценном многоканальном звуке, который переносит слушателя в заданное акустическое пространство с множеством локализованных источников. Но, увы, это пока не реально.

Пока, можно говорить только об одной успешной реализации, наиболее близкой к нашим реальным ощущениям - бинауральной. Для получения полноценной объемной информации с ее помощью слушатель должен находиться в наушниках. Опытная модель, на базе которой строились основные исследования, представляла собой манекен, где вместо правого и левого уха были установлены микрофоны. По результатам составлялась сложная математическая модель, которая позволила создавать виртуальные звуковые миры.

Но эти исследования, хоть и были близки к идеалу, имели достаточно общие результаты. Ведь, согласитесь, у каждого человека различная форма ушных раковин, а расстояние между правым и левым ухом - это также индивидуальный параметр. В общем, исследователи зашли в тупик. Практическое применение результатов бинауральных моделей сейчас предлагается в варианте субкультуры, развивающейся отдельно от основных направлений, применяемых на практике.

Как это ни странно, но самый важный вклад в области пространственного звука сделала киноиндустрия. В процессе своего развития она сталкивалась с множеством проблем, и нельзя сказать, что ее путь был гладким. На внедрение новых технологий тратились огромные деньги, и не всегда такие вложения оправдывались. В качестве примера можно привести стереокино, где зрителям предлагалось надевать специальные очки, чтобы смотреть изображение в объеме. Стереокино не стало стандартом, и даже хуже - оно полностью не прижилось. Поэтому идеализированные бинауральные варианты с наушниками - технология не для кинопроизводства.

### **Dolby**



Как вариант - это многоканальная система. Изначально кино было немым, потом пришел моно звук, и следующий этап - стерео... Стерео ворвалось с большим количеством новых идей и их реализаций. Простой поворот ручки панорамы приводил зрителей и слушателей в восторг. Что интересно, до 80-х понятия "stereo" и "surround" являлись синонимичными. А знаете, когда разделились? С переходом стандарта киноплёнки с 35 мм на 70 мм. Изначально, производители стали использовать так называемый канал эффектов (effect channel) - это расположенный позади слушателей громкоговоритель, предназначенный для озвучивания вспомогательных звуков, а зачастую и просто молчащий, поскольку место на 35 мм-плёнке было ограничено. С внедрением стандарта 70 мм, канал эффектов стал задействоваться постоянно и был переименован в пространственный канал (surround channel, далее будем называть его как S-канал). И с этого момента можно сказать о разделении понятий "stereo" и "surround". В варианте 35 мм-плёнки было предложен стандарт Dolby Stereo (середина 70-х). В 1982 году на рынок поступили первые Dolby Surround устройства. Стоили они тогда как аппаратура hi-end-класса, и были доступны очень небольшому числу граждан.

Параллельно с киноиндустрией развивались видео-индустрия и кабельное/спутниковое телевидение. Возможность обеспечить объемный звук в домашних условиях - одна из основных задач, которая была поставлена перед производителями. Появились более адекватные как по цене, так и по реализации варианты декодеров Dolby Pro Logic и Pro Logic II. Последний позволял не только воспроизвести закодированный объемный звук, но и слушать музыку с обычного CD, раскладывая его на surround-систему. Pro Logic и Pro Logic II являются переходными мостиками между обычными стерео-системами и многоканальным surround. Утрируя некоторые особенности можно сказать, что в чем-то они (стерео и surround в вариантах декодеров Dolby Surround - Pro Logic и Pro Logic II) взаимозаменяемы.

С активным внедрением цифровых технологий Dolby Labs представляет стандарт Dolby Digital AC-3 (5.1), разработанный для 35 мм пленки. Особенностью Dolby Digital 5.1 является наличие двух surround-каналов (правого (RS) и левого (LS), у предшественников S-канал был один, хотя громкоговорителей устанавливалось несколько) и сабвуфером. Если говорить об аббревиатуре "5.1", то она расшифровывается не только как пять каналов и сабвуфер. Дело в том, что в Dolby Digital впервые фронтальные и S-каналы имели одинаковый полный частотный диапазон. Например, в Dolby Pro Logic S-канал предусматривал диапазон от 100 Гц до 7КГц. Теперь же мы говорим о пяти полноценных каналах с диапазоном 20 Гц - 20 КГц и выше. Частотный диапазон сабвуфера или как его еще называют LFE-канал (Low Frequency Effect) является ограниченным и составляет примерно 0.1 от частотного диапазона пяти каналов, поэтому "5.1" следует расшифровывать как "пять и одна десятая".

Файлы Dolby Digital 5.1 имеют расширение \*.ac3 и требуют для своей расшифровки специальный декодер (программный или аппаратный). На заре своего становления технология стоила очень не дешево. Декодеры были только аппаратные и выпускались в виде специализированных устройств. Теперь ситуация изменилась в лучшую сторону, и обычный пользователь может себе позволить приобретение звуковой карты со встроенными декодерами или скачать их программный вариант из интернета. Таким образом, современный кинозал и домашний кинотеатр полностью уравнились по возможностям. Стоит отметить и еще один немаловажный аспект.

Наряду с развитием кино, видео и телевидения, открылся новый спектр применения технологии surround-звука - видео и компьютерные игры. Если раньше мы говорили о простейших вариантах 3D-движка, основанном на регулировках фазовых сдвигов, задержек и простейшей эквализации, то теперь подразумеваем полноценный 5.1 как минимум.

На сегодня Dolby Digital 5.1 является самым универсальным стандартом. Звук AC-3 может быть преобразован в зависимости от масштабов аппаратуры. Например, помимо доступного разделения на пять или шесть каналов (в зависимости от наличия/отсутствия сабвуфера), он может преобразовываться в четыре канала (Dolby Pro Logic), обычно стерео и даже моно. Поэтому, кодируя звук для компьютерных игр в Dolby Digital 5.1 можно не бояться того, что

пользователь, имеющий в наличии меньшее количество колонок, будет чувствовать себя неудобно.

Естественно, кинотеатрам нужно иметь преимущество по сравнению с домашними комплексами, поэтому специально для них была разработана технология Dolby EX, подразумевающая ничто иное как 6.1. Помимо отдельных S-каналов для правого и левого (RS и LS) в ней предусмотрен S-канал центрального (CS). Технология Dolby Digital позволяет компрессировать данные и передавать их потоком с минимальной скоростью передачи 320 kbps (стандарт - 384 kbps). Конечно, сжатые данные являются более транспортабельными и занимают меньше места на носителях, будь то киноплёнка или обычный винчестер. Но данная технология имеет и минусы - она рассчитана только на один цикл кодирования-декодирования и не решает проблемы ретрансляции и редактирования сжатых данных. Помимо этого хоть мы и выигрываем в месте, мы проигрываем в том, что звук стандарта AC-3 не синхронизирован с видеоданными, поэтому редактирование такого материала весьма проблематично.

Естественно, требуется многократное кодирование/декодирование. Решением этой проблемы явилась новая внедренная технология Dolby-E. Она позволяет передавать, многократно кодировать/декодировать восемь каналов с полным частотным диапазоном. Всего допускается десять циклов кодирования/декодирования. Помимо всего прочего стандарт предусматривает наличие метаданных - специальных инструкций, описывающих основные данные. Кроме синхронизации они могут нам сообщить настройки компрессора для каждого канала и т.п. Кстати, этими данными можно и не пользоваться, они не больше ни меньше, а дополнительная информация..

## DTS



У основной массы пользователей понятие surround ассоциируется с Dolby. Это не так. Dolby Digital - более раскрученный брэнд, не более. В середине 80-х Терри Бирд, президент и основатель фирмы Nuoptix, объединился в команду с Джимом Кетчамом. Задачи, которые поставили перед собой экспериментаторы, были следующими - создать единый цифровой звуковой стандарт для кинотеатров и решить вопросы качественного размещения многоканального аудио на киноплёнке. Изначально эксперименты велись с помощью цифровых магнитофонов (DAT) и проектора. Уже на начальном этапе разработки

решались вопросы синхронизации (как мы знаем, в варианте Dolby они разрешились только с внедрением Dolby-E). Через некоторое время вместо DAT стали использоваться обычные CD-ROM. При использовании алгоритма сжатия apt-X100, разработанного Стивеном Смитом и его коллегами из Audio Processing Technology, позволяющего компрессировать данные 4:1, на один CD-ROM помещалось 100 минут многоканального аудио (скорость передачи данных 884 kbps). Таким образом на двух дисках мог храниться звук для трехчасового художественного фильма.

После получения патента в 1990-м имела демонстрация новой технологии для членов SMPTE. К 1992-му году технология была показана Стивену Спилбергу, правда, на базе винчестеров, и режиссер согласился ее применить в своей новой картине "Парк Юрского периода" ("Jurassic Park"). Студия Universal, являющаяся хозяином картины, решила проверить технологию на нескольких двух "менее значимых" фильмах. Результаты тестов оказались приемлемыми. Более того - Universal и Стивен Спилберг стали совладельцами с Терри Бирдом новой фирмы получившей название Digital Theatre Systems. Только за то время, когда образовывалась эта компания (а это заняло четыре месяца), в кинотеатрах было инсталлировано 876 систем DTS. На сегодняшний момент оных более 19.000 по всему миру. Было запланировано реализовать два технологических стандарта:

**DTS-S** - DTS-стерео, предназначен для кинотеатров, которые хотят использовать цифровой звук, но не могут себе позволить полноценный шестиканальный вариант.

**DTS-6** - три фронтальных канала (левый L, центральный C, правый R), два S-канала (LS и RS) и сабвуфер. При этом следует отметить, что стандарты несколько различны с Dolby. Например частотный диапазон сабвуфера в DTS - от 20 до 80 Гц, вместо 3 -120 Гц у Dolby Digital.

Одним из главных отличий DTS от Dolby Digital является то, что DTS разрабатывался для 70 мм киноплёнки. Соответственно, проблема места не стояла так актуально как в варианте Dolby Digital, вопросы реализации hi-res-звука и синхронизации решались практически сразу со времени возникновения стандарта.

Сейчас, упоминая DTS, мы можем говорить о восьми и более каналах hi-res-звука, новейших технологиях сжатия информации, и их активном практическом применении.

### **Dolby vs DTS**

Стандарты DTS распространены точно также как и Dolby. При этом, конкуренты успевают ссориться. А что им еще делать? У Dolby гораздо мощнее поддержка со стороны обычных потребителей и производителей технологий для пользовательского рынка. Так, в Dolby Surround выпускаются музыкальные альбомы звукозаписывающими лейблами Delos, RCA Victor/BMG Classic, Concord Jazz. Документация к этому стандарту постоянно востребована разработчиками современных видеоигр. Современное спутниковое телевидение также вещает в большинстве своем со звуком в режиме Dolby Surround. На

последнем 114-м съезде AES в Амстердаме имели место обучающие семинары по записи, сведению и мастерингу в Dolby 4.0, 5.0, 5.1, 6.0, 6.1, и 10.2.

DTS остается привилегией кинопроизводства. На сегодня уже успешно решен вопрос с hi-res-аудио с фантастическими на сегодняшний день стандартами 64 бита и выше, 192 КГц и выше. Кстати, Dolby Digital также позволяет использовать hi-res-треки.

Производителями ПО и студийной техники вопрос противостояния решен очень просто - поддерживается и то, и другое. Это можно увидеть и в Steinberg Nuendo - одном из самых профессиональных и универсальных средств для производства surround-звука, и в линейке звуковых интерфейсов Audiotek - MAYA 5.1, 7.1, EX, EX7, PRODIGY 7.1. Там вы встретите поддержку сразу двух технологий - Dolby Digital и DTS. Помимо этого, стандартные программные приложения, предназначенные для воспроизведения звука в DVD также имеют декодеры для обоих стандартов.

Если же посмотреть на конфронтацию DTS и Dolby, то иногда доходит до смешного. Dolby выпускает некий пресс-релиз - DTS отвечает. Например: "Dolby проводила исследования на неизвестной студии с неизвестной аппаратурой при этом для воспроизведения использовался неизвестный материал, который тестировали неизвестные специалисты...". Да и иногда, говоря о surround звуке и DVD, Dolby действительно забывается, присваивая основные свершения и т.п. в свою пользу. В общем, обычная корпоративная война.

## **7-Лекция. Другие многоканальные системы**

В принципе, любой пользователей, имея в своем арсенале звуковую карту с четырьмя и более выходами и профессиональную мультитрековую программу, может почувствовать себя на заре 70-80-х. Именно тогда велись основополагающие опыты с многоканальным звуком. Многим известна квадрофония, была даже октофония. При этом, например, октофония тогда не подразумевала современных "7.1". Тогда предполагалось наличие 8 независимых каналов с полным частотным диапазоном, расставленных вокруг слушателя по кругу на равных расстояниях между собой.

Экспериментировать можно сколько угодно. Как результат проводимых исследований -необходимость стандартизации в сведении и воспроизведении. Для примера, возьмем вариант октофонии, о котором мы вспомнили чуть выше.

Исходные данные - ударная установка, записанная на 14 микрофонов и сведенная таким образом, чтобы слушатель, находящийся в центре получал ощущение присутствия. При изменении положения слушателя от центра, близлежащие к нему громкоговорители будут громче, баланс и звуковая картина изменятся. Таким образом, в одной точке мы слышим одно, в другой - другое. Это основной недостаток систем такого плана - слушатель должен располагаться в центре.

В варианте surround нам сообщается только пространственный эффект присутствия. То есть локализация звука закреплена за отдельным каналом. Другие работают как сопровождающие. То есть, мы получаем эмуляцию пространственных ощущений, для чего собственно surround-системы и разработаны.

Самая уникальная многоканальная акустическая система - это наушники, в которых можно получить качественный 3D.

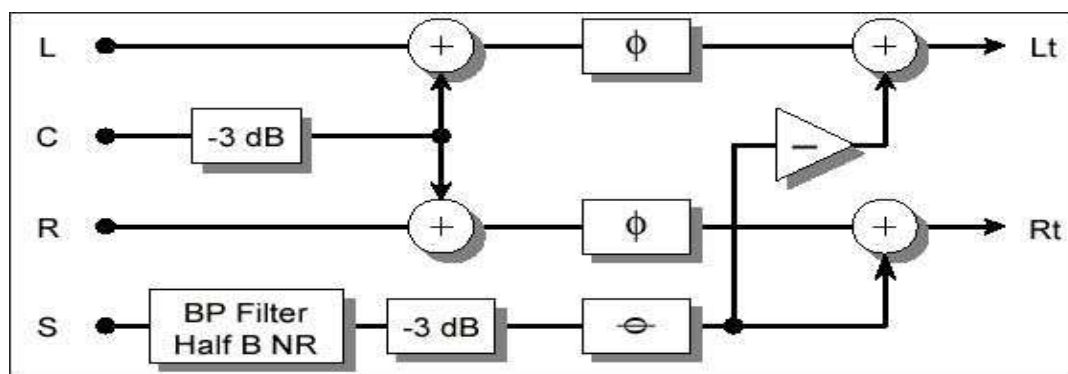
Хотя с другой точки зрения, никогда не стоит останавливать эксперименты. У вас есть к этому все возможности. Если в середине 90-х для опытов с квадрофонией мне приходилось подключать в PC две стерео-платы, то теперь даже самый дешевый звуковой интерфейс имеет как минимум четыре линейных выхода. Dolby Surround не является идеальной системой и места для экспериментов очень много.

### Стандартный кодек Dolby Surround и пассивный декодер

Давайте немного вернемся к истории, в начало 80-х, когда появились первые Dolby Surround устройства. Кодек Dolby Surround - это устройство, позволяющее преобразовать четыре канала (L-левый, C-центральный, R-правый и S-пространственный) в два кодированных Lt и Rt. Lt и Rt - это аббревиатура от L-, R- total, что в переводе означает общий. В документациях к Dolby Surround такое наименование обозначает закодированные каналы для систем Dolby старого поколения. Такие Lt и Rt могут быть записаны на обычную магнитную пленку и воспроизводиться на ленточных магнитофонах. При наличии декодера, Lt и Rt расшифровываются на многоканальную систему Surround.

Давайте рассмотрим работу кодера Dolby Surround на представленной схеме. На рис. 1 отображено как четырехканальная система преобразуется в двухканальную. Обозначения и названия для каналов являются стандартными для всех систем Dolby Surround:

- L** - фронтальный левый канал.
- C** - центральный канал.
- R** - фронтальный правый канал.
- S** - surround-канал.
- Lt** - общий левый.
- Rt** - общий правый



Принципиальная схема работы кодера Dolby Surround



Как мы видим в процессе преобразования, сигнал центрального канала делится поровну и микшируется с левым и правым, при этом его уровень понижается на 3 Дб. Сигнал S проходит обработку через отдельный тракт, в состав которого входят: bandpass-фильтр - ограничение частотной полосы в диапазоне 100 Гц - 7 КГц (surround-канал имеет частотный диапазон 100 Гц - 7 КГц), шумоподавление Dolby B-type Noise Reduction, коррективировка фазы - составляющие сигнала S сдвигаются по фазе таким образом, чтобы при сложении с левым и правым каналом оказаться в противофазе.

В варианте пассивного декодера сигнал S-канала можно получить простым вычитанием Lt-Rt. На рис.2 представлена схема работы простейшего декодера Dolby Surround, на выходе которого мы имеем L, R и S.

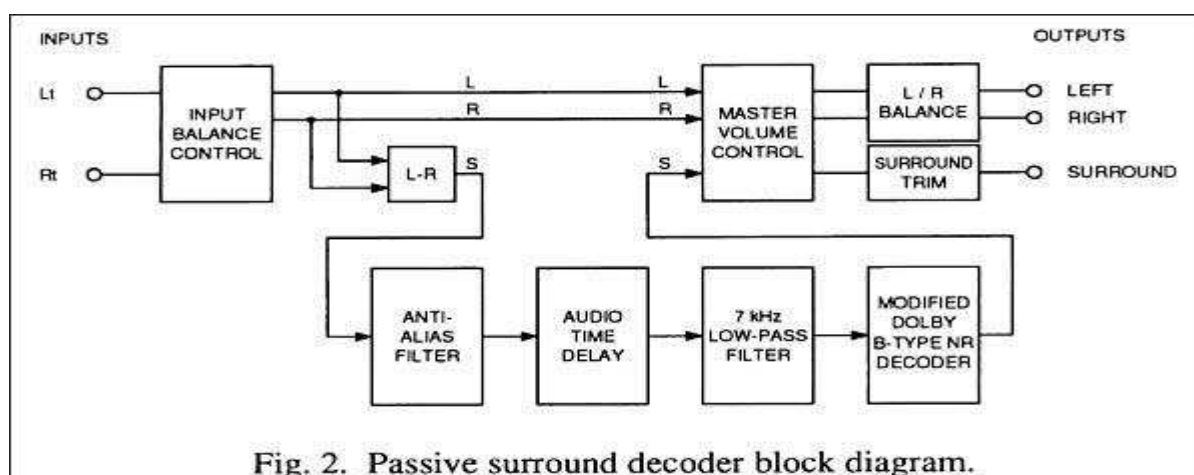


Fig. 2. Passive surround decoder block diagram.

При декодировании возникает несколько проблем, главная из которых - взаимопроникновение фронтальных каналов и S. Например, при изменении равновесия уровней Lt и Rt в сигнал S-канала проникают составляющие и L, и R. Этот вопрос решается двумя путями - задержкой устанавливаемой на сигнал S, и ограничением его частотного диапазона. При задержке, равной 10 мс, устраняется эффект Хааса - сигнал с фронтальных каналов идет с опережением и локализация источника ассоциируется с ними. Ограничение звука до 7КГц также позволяет акцентировать внимание на фронтальных каналах, поскольку их звучание ярче, чем у S-каналов.

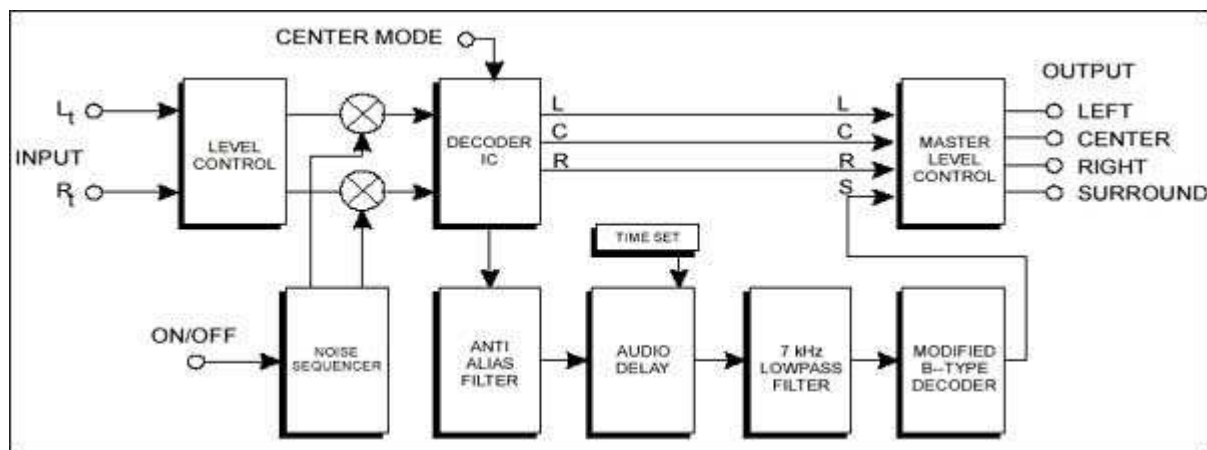
В таком варианте пассивного декодирования подразумевается, что центральный канал является виртуальным и соответствует панорамному центру стереосистемы L и R. Данная система идеализирована, поскольку громкоговорители правого и левого каналов должны быть равноудалены от слушателя, что на практике практически недостижимо (в кинотеатре много слушателей).

### Активные декодеры Dolby Surround. Pro Logic

Активные декодеры Dolby Surround - это новое поколение преобразующих устройств от Dolby Labs, внедренное в середине 80-х. Используя в качестве исходных данных те же Lt и Rt, что и в варианте пассивного декодера, они предусматривают более сложный процесс декодирования с учетом наличия



физического (а не виртуального центрального канала). Первым устройством подобного класса стал Dolby Prologic.



### Декодер Dolby Pro Logic.

Условно схему работы Pro Logic можно разделить на две составляющие - пассивную и активную. В пассивной части сигнал S-канала получается простым вычитанием сигналов левого канала из правого, а сигнал центрального канала получается путем сложения правого и левого. Как мы знаем при кодировании S-канал записан в противофазе и при таком сложении он должен исчезнуть. Чтобы центральный канал был слышимым, сигналы каналов L и R ослабляются на 3 Дб. Другой вариант - удаление сигнала центрального канала из левого и правого каналов. Делается это также с помощью противофазного сложения и взаимной компенсации компонентов сигналов правого и левого каналов. Данный принцип называется Cancellation Concept (принцип взаимной компенсации).

Если в процессе кодирования и записи было изменено балансное соотношение между Lt и Rt, то это может повлечь за собой ряд проблем - в сигнал S-канала проникают составляющие и L, и R, и так далее. В варианте Pro Logic для выравнивания баланса между входящими Lt и Rt предусмотрен специальный модуль Level Control.

Активная часть данного декодера позволяет выделять звуковые образы. Данная система предназначена только для кино. Параллельно со звуком идет и видео. Соответственно, основной задачей Pro Logic является обеспечение ассоциативного ряда. В описании к Pro Logic существует термин "доминирующий звуковой образ" (dominant sound). Появление этого термина не случайно и имеет прямое отношение к психоакустике. При акцентировании внимания зрителя на доминирующем звуковом образе все остальные звуковые объекты "размазываются" в пространстве и не имеют четкой локализации. В варианте Dolby это называется Signal Masking Concept (принцип маскирования сигнала). Именно он и является основополагающим для систем активных декодеров. Соответственно, в этих устройствах сигнал речи являются управляющим, все остальные - управляемые. Если киногерои говорят и параллельно с этим идет стереомузыка, то их речь воспроизводится центральным каналом, усиление L и R уменьшается. Музыка воспроизводится

и за счет громкоговорителей S-канала. При замолкании речи звуковая картина и баланс уровней каналов восстанавливаются и приходят в исходное состояние. Таким образом, мы получаем некое "качание", называемое pumping.

Если персонажей несколько, их речь отличается от других звуков по мощности, и они локализованы в разных местах, то активный декодер Pro Logic работает в так называемом "быстром" режиме, переключая настройки с одного доминирующего образа на другой, время реакции декодирующей системы минимально. Если же звуки не имеют значительных расхождений по мощности, то декодер переходит в "медленный" режим.

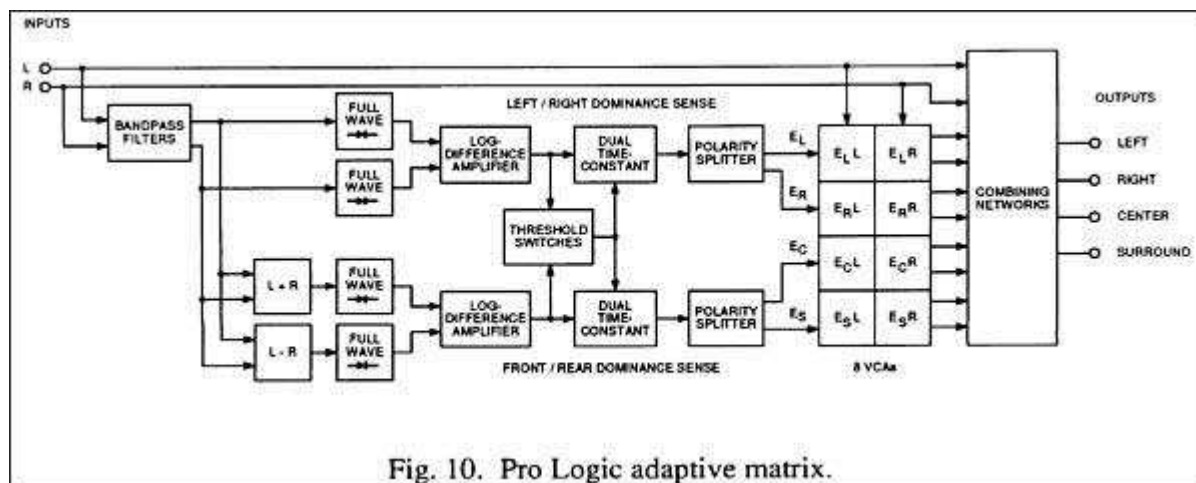


Fig. 10. Pro Logic adaptive matrix.

Адаптивная матрица Dolby Pro Logic.

Говоря о системах Pro Logic, стоит упомянуть немаловажный аспект, который их выгодно отличает от предшественников - декодер может определять направление локализаций звуковых образов. Достаточно взять амплитуду сигнала центрального канала и сравнить ее со значениями амплитуды этого же сигнала в правом и левом каналах, и мы можем определить вектор направления локализации звукового образа. Дальше - дело техники.

Сердце Pro Logic - это адаптивная матрица. Для анализа звуковой картины используется четыре входящих сигнала - Lt, Rt, Lt+Rt и Lt-Rt подаваемые на два дифференциальных усилителя (на первый - Lt и Rt, на второй - Lt+Rt и Lt-Rt). В результате на выходах этих усилителей выделяются два управляющих сигнала, являющиеся ничем иным как разностями логарифмов от амплитуд поступающих сигналов. Не смотря на то, что в результате получается всего два управляющих сигнала, они являются биполярными, что позволяет устройству произвести полный анализ происходящего. При этом в данной матрице предусмотрено сравнение амплитуд с пороговыми значениями, что позволяет переключать "быстрый" и "медленный" режимы.

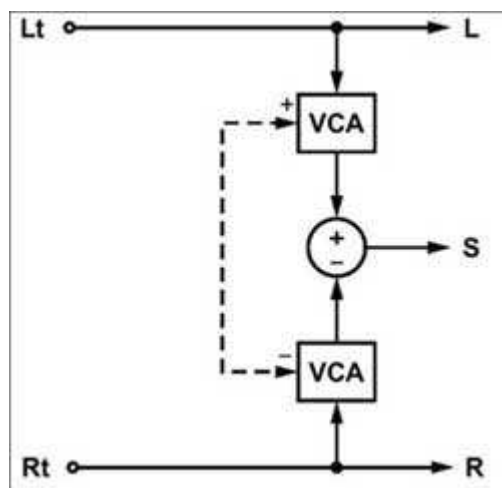
На выходе адаптивной матрицы мы имеем блок анализа, который оперирует десятью сигналами, складывая и вычитая их в определенной весовой пропорции. В результате мы имеем четыре стандартных сигнала C, R, L и S.

## Pro Logic II

Dolby Surround Pro Logic II - это следующий этап развития аналоговых систем декодирования и последний из них. Он намного проще по принципу работы по сравнению с Pro Logic, помимо этого позволяет производить

прослушивание в режиме surround не только звуковых треков к кинофильмам, но и музыкальных записей Dolby Surround. Начиная с этого этапа можно говорить об удешевлении стандартов Dolby, их приходе на рынок бытовых устройств.

Чтобы понять простоту работы Pro Logic II рассмотрим две схемы. Первая - реализация формирования сигнала S-канала.



Формирование сигнала S-канала

Как мы уже говорили в описании к Pro Logic, очень большое значение имеет отсутствие проникновения сигналов  $L$  и  $R$  каналов в  $S$ -канал. Но, например, если баланс между каналами изменен и доминирующий образ расположен в направлении, к примеру, "центр-право", то такое проникновение будет обязательно. На рис. 5 показан вариант простейшего решения данной проблемы. В цепи между входами  $L_t$  и  $R_t$  и входом сумматора ставятся два управляемых усилителя VCA (Voltage Controlled Amplifier). Управление обоими VCA взаимосвязано в разной полярности. Если сигнал  $L_t$  мощнее  $R_t$ , то усиление сигнала  $L_t$  перед сумматором уменьшается, а  $R_t$ , наоборот, увеличивается. Таким образом, сигналы выравниваются - проникновение минимально.

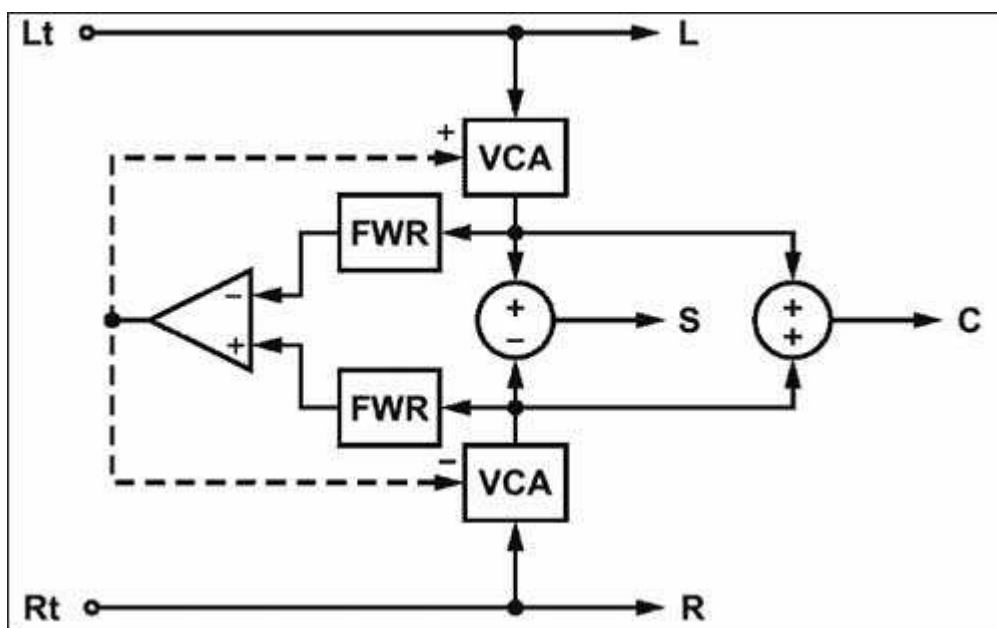


Схема работы Pro Logic II.

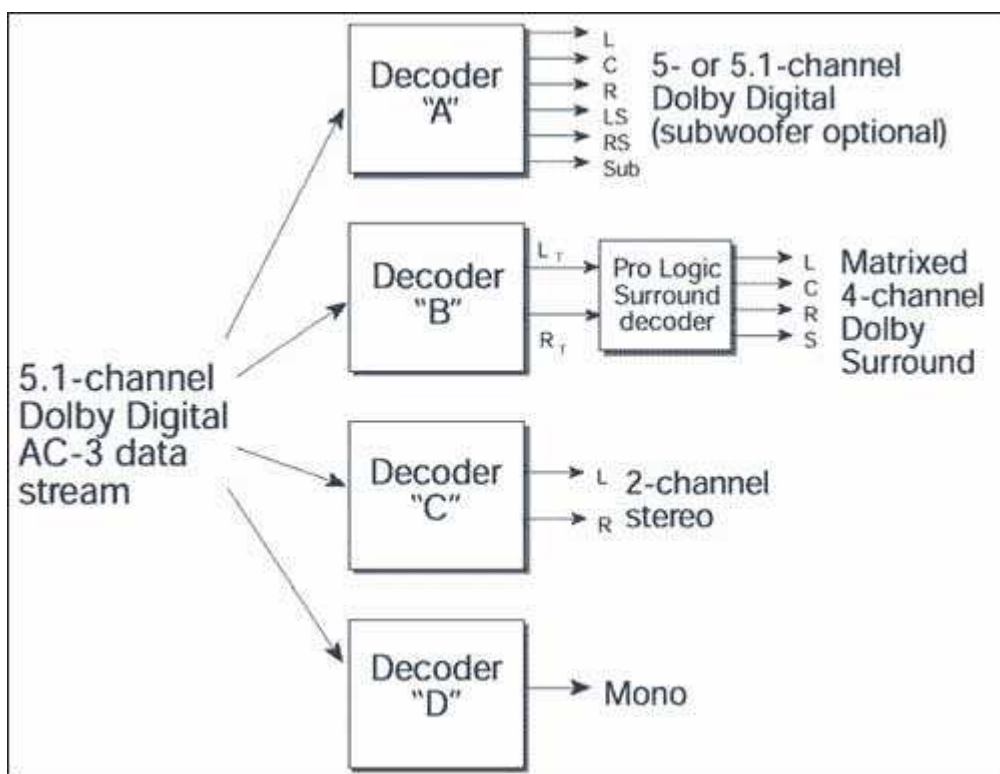
Схема работы Pro Logic II также проста в понимании. Сигналы с обоих VCA поступают на модули амплитудных детекторов FWR (Full-Wave Rectifier). Дифференциальный усилитель сравнивает сигналы постоянного тока, которые пропорциональны амплитудам сигналов Lt и Rt. Его сигнал является управляющим для VCA. Сигнал центрального канала С получается путем сложения выровненных Lt и Rt.

По сравнению с Pro Logic, Pro Logic II не имеет "быстрого" и "медленного" режимов - переменная времени реагирования изменяется постоянно. Схема его работы проста и удобна.

### **Dolby Digital AC-3**

Dolby Digital AC-3 или как его еще называют Dolby Digital 5.1 появился в конце 80-х и был разработан для 35-мм пленки. С приходом этого стандарта нас ожидало несколько новшеств. Во-первых, был разработан шестиканальный стандарт, в котором появилась возможность использования сабвуфера, или как его еще называют LFE-channel (Low Frequency Effect). Во-вторых, S-канал перестал быть моно. То есть, в системе 5.1 мы имеем в наличии SL и SR-каналы (Surround Left и Surround Right). Громкоговорители всех каналов кроме LFE являются широкополосными. Впервые стандарт предложил возможности сжатия динамического диапазона. При этом стандарт AC-3 предусматривает компрессию данных. Сейчас скорость цифрового потока аудиоданных Dolby Digital составляет 320 кбит/сек и больше.

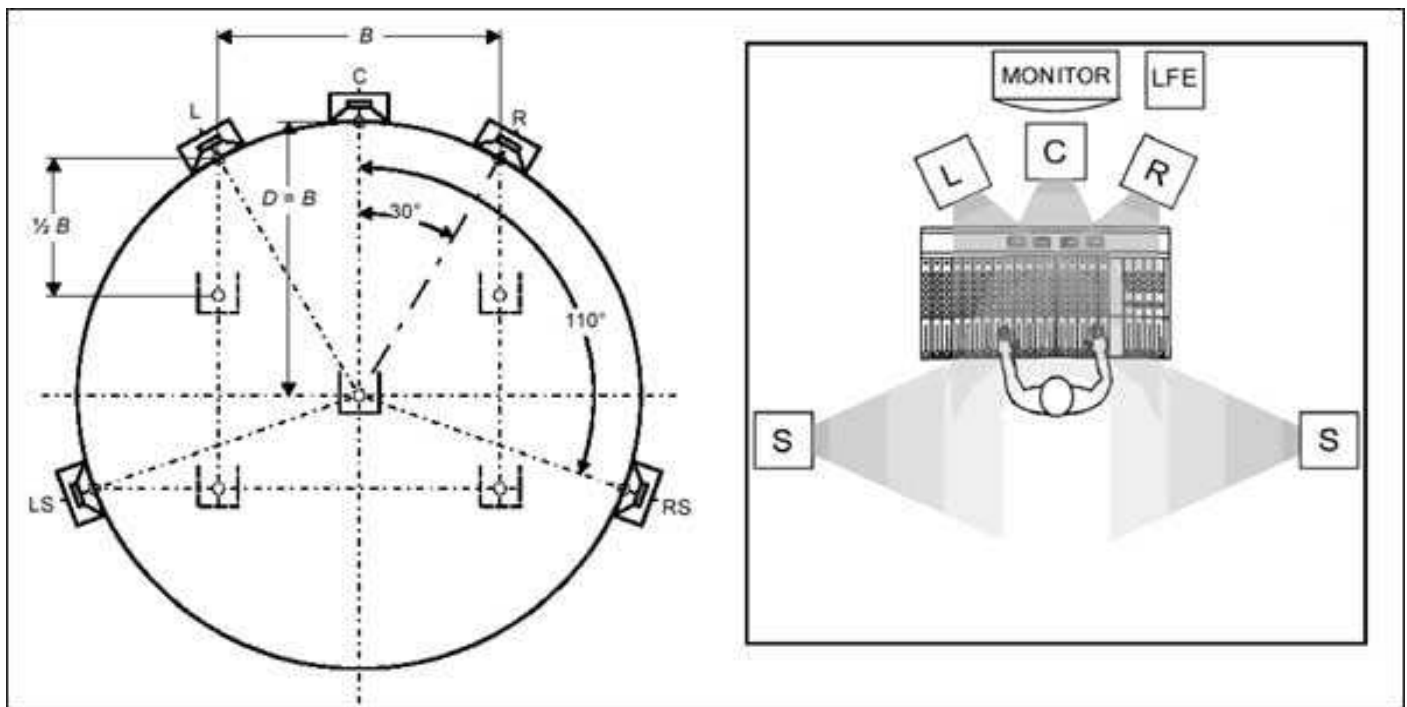
Главная особенность Dolby Digital AC-3 - это его масштабируемость.



#### Масштабирование Dolby Surround 5.1.

Например, мы написали музыку в стандарте 5.1. При этом пользователь может ее прослушать без потерь в качестве в 5.1 (5.0), четырехканальном варианте Pro Logic, в обыкновенном стерео и даже в моно. И теперь представьте как это выгодно и пользователям и производителям. Компьютерные и видео-игры нового поколения само собой подразумевают наличие звука 5.1. При этом, пользователи, не имеющие расширенных вариантов акустических систем не будут чувствовать себя дискомфортно - система будет масштабироваться под параметры системы, имеющейся в наличии. Многие спутниковые телеканалы вещают в 5.1, но это ни как не отражается на большинстве зрителей, имеющих телевизоры с моно-звуком.

## 8-Лекция. Настройки многоканальной системы



Правильная расстановка акустических систем для Dolby 5.1  
Рис. 8

### Что такое THX?

За часто встречающейся аббревиатурой THX стоит целый набор технологий и стандартов для современных многоканальных систем. Этот сертификат был введен Джорджем Лукасом и его фирмой Lucas Films, и расшифровывают эту аббревиатуру по разному. Первая версия - "THX" была названа в честь первого фильма Лукаса - "THX-1138", вторая - "THX" как аббревиатура от Tomlinson Holman eXperiment.

Основные требования THX для кинотеатров выглядели следующим образом:

- отличная разборчивость экранного диалога;
- передача всего динамического диапазона фонограммы, от легкого шепота до самого громкого взрыва без искажений;
- однородное озвучивание всех зрительских мест;
- ровной и широкой частотной характеристикой (бас с 30 Гц с одинаковым уровнем громкости во всем зале).

После того, как с кинотеатрами было покончено (сейчас такие лицензии имеют не более чем 900 кинотеатров по всему миру), команда стандартизации обратила внимание и на домашние устройства. По сути, THX является ничем иным как "Джордж Лукас рекомендует...". Это похоже на то как Арнольд Шварценеггер рекламировал мультимедийную продукцию ВВК. Но вариант Лукаса более выигрышный, поскольку он привлек специалистов, которые

создали вполне понятные технические нормативы, во-вторых, он сам поднял этот вопрос.

### **THX Ultra**

THX Ultra распространяется на акустические системы в помещениях с объемом свыше 86 кубических метров (3000 кубических футов). Стандарт включает в себя:

*Cinema Re-EQ.* Система реэквализации, устраняющая излишнюю яркость звука, появляющуюся при воспроизведении в домашних условиях звуковых дорожек, рассчитанных на кинозалы.

*Timbre Matching.* Тональное пересогласование фронтальных и пространственных каналов.

*Adaptive Decorrelation.* Обеспечение стерео-ощущений при наличии моно звука в пространственных каналах.

*Bass Management Electronic Crossover.* Обеспечение качественного прослушивания низких звуков через сабвуфер применительно к домашним системам.

*Bass Peak Level Management.* Защита сабвуфера от перегрузки.

*Loudspeaker Position Time Synchronization.* Настройка акустического пространства под месторасположение зрителя/слушателя.

### **THX Select**

THX Select распространяется на акустические системы в помещениях с объемом около 57 кубических метров (2000 кубических футов). Уровень в 85 дБ позволяет выполнить основное условие стандарта - прослушивание самых тихих и самых громких звуков. Этот уровень часто называют референсным. В контроллерах и ресиверах установлены специальные цифровые индикаторы, отображающие разницу между текущим и референсным значениями уровней.

### **THX Multimedia**

Вот тут самый интересный момент, так как его считают одним из самых ангажированных. THX Multimedia был разработан в 1999 году для сертификации мультимедийных компьютерных систем. Смотрим небольшой список счастливиц...

Сертифицированные колонки THX Multimedia:

Altec Lansing ADA 885 4.1 System

Cambridge SoundWorks® THX® 5.1 550 speaker systems

Creative MegaWorks® THX® 5.1 550 speaker systems

Klipsch ProMedia 2.1 System

Klipsch ProMedia 4.1 System

Klipsch ProMedia 5.1 System

Logitech Z-560 4.1 System

Сертифицированные системы THX Multimedia:

Dell Dimension 4400 System

Dell Dimension 8200 System

Сертифицированные аппаратные звуковые интерфейсы THX Multimedia:

Creative Labs' Sound Blaster Audigy 2

## **9-Лекция. Компьютерные программы для редактирования звуковых файлов.**

Развитие компьютерных технологий сделало доступным почти каждому такие виды творческой деятельности, которые еще лет двадцать назад были уделом профессионалов. Сейчас сочинять, записывать и редактировать музыку может любой владелец современного персонального компьютера.

Конечно, все имеет свои плюсы и минусы, и доступность компьютерных технологий привела к тому, что хорошая музыка рискует оказаться погребенной под горами дилетантского мусора. Разумный человек, хоть раз попробовавший себя в роли композитора, быстро понимает, что сочиняет музыку не компьютер, а человек с талантом, который еще никому не удалось формализовать в программном коде. Но вот научиться редактировать аудиофайлы, в том числе и музыкальные, вполне по силам каждому, кто снимает любительские видеофильмы, коллекционирует музыку или реставрирует старые, но ценные записи.

Для решения этих и других задач используют программы, которые называют аудиоредакторами. Их существует довольно много, и спорить о том, какая лучше, можно бесконечно, точно так же, как автомобилисты до хрипоты спорят о достоинствах "японок" по сравнению с "немками". Наиболее известными аудиоредакторами считаются Sound Forge ("Кузница звуков"), Cool Edit Pro (ныне Adobe Audition) и WaveLab. Каждая из этих программ имеет свои достоинства и недостатки, использует фирменные технологии обработки звука, но в целом задачи они решают одинаковыми способами, да и графический интерфейс у них схож.

Компьютерный звук может быть преподнесён двумя типами файлов - звуковыми и музыкальными.

Первые представляют собой записанную с помощью микрофона "живую" музыку, человеческую речь, пение птиц, шум леса, словом, что угодно.

К музыкальным файлам (MIDI-файлам) относят композиции, созданные на компьютере, то есть не имеющие природного аналога - фактически это нотная запись. Аббревиатура MIDI происходит от слов Musical Instrument Digital Interface, это интерфейс для связи синтезаторов или других электронных музыкальных инструментов с компьютером или друг с другом.

Аудиоредакторы имеют некоторое сходство с графическими программами. Те, что предназначены для звуковых файлов, напоминают программы для растровой графики, а секвенсоры, служащие для создания и редактирования музыкальных файлов - программы для графики векторной.



Подобно растровым изображениям, звуковые файлы занимают большой объем, их сравнительно сложно редактировать, но зато и звук получается богатый, живой, насыщенный. Музыкальные файлы компактны, их намного проще создавать, редактировать, преобразовывать, но и результат гораздо беднее. Полноценную музыку с помощью секвенсоров "сочинять" очень трудно.

На самом деле, конечно, между композициями на основе звуковых и музыкальных файлов нет каменной стены, нередко основная часть произведения записывается в виде звукового файла, а часть музыкальных партий пишется по MIDI. MIDI-музыку вообще можно собирать из готовых кусочков, используя заготовки-петли, то есть зацикленные последовательности звуков, например барабанные "кольца". Современные танцевальные вещи часто так и создаются. Но в этой статье мы будем говорить только о редактировании звуковых файлов.

Современные редакторы звука, работающие под управлением MS Windows, в обязательном порядке используют программный интерфейс DirectX. Первоначально он был предназначен для повышения быстродействия Windows при обработке больших объемов изображения и звука.

DirectX - интерфейс низкого уровня. С помощью своих API он предоставляет программисту прямой доступ к памяти графических и звуковых адаптеров, за счет чего, собственно, и достигается ускорение работы программ. DirectX состоит из нескольких модулей, основными из которых являются: DirectDraw - модуль двумерной графики; Direct3D - модуль трехмерной графики; DirectSound предназначен для работы со звуком; DirectInput - для клавиатуры, мыши, джойстика и аналогичных устройств; DirectPlay - в основном для поддержки сетевых игр; DirectAnimation - для анимационных эффектов на WEB-страницах, а DirectMusic - для поддержки музыки в компьютерных играх.

В современные операционные системы типа Windows 2000 и Windows XP DirectX уже встроен, и удалить его обычными средствами не получится, его можно только обновлять по мере выхода очередных версий. На момент написания этой статьи самой новой является версия DirectX 9.0с, скачать которую для Windows 98, ME, 2000 и XP можно с сайта Microsoft по следующей

ссылке:  
<http://www.microsoft.com/windows/directx/default.aspx?url=/windows/directx/downloads/default.htm>

Хорошие аудиоредакторы имеют открытую архитектуру и могут работать со множеством плагинов, разрабатываемых сторонними фирмами. Плагины чрезвычайно расширяют возможности программы, что еще раз подчеркивает их сходство с редакторами растровой графики, только плагины к Adobe Photoshop называются фильтрами. И сами аудиоредакторы, и плагины к ним активно

используют DirectX, так что на старую операционную систему с устаревшей версией DirectX современный аудиоредактор, скорее всего, установить не удастся.

Все операции, которые способны выполнять аудиоредакторы, условно можно разделить на четыре группы: простейшее редактирование, звуковые процессы, звуковые эффекты и дополнительные инструменты.

Операции простейшего редактирования не затрагивают внутренней структуры звука, то есть не изменяют его акустических и частотных характеристик. К операциям простейшего редактирования обычно относят такие, как: Cut (Вырезать выделенное), Clear/Delete (Очистить/Удалить), Trim/Crop (Вырезать все, кроме выделенного), Copy (Копировать), Paste (Вставить), Mix (Смешать), Crossfade (Плавно заместить). Для тех, кто имеет хотя бы минимальный опыт работы с компьютером, смысл этих операций очевиден, за исключением разве что двух последних.

При выполнении операции Mix (Смешать) содержимое буфера обмена смешивается с данными в открытом окне, начиная с позиции курсора, а при выполнении Crossfade (Плавно заместить) данные в окне плавно замещаются содержимым буфера, начиная с позиции курсора, причем громкость исходного сигнала плавно уменьшается, а громкость другого - плавно нарастает.

Вообще при выполнении всех операций простейшего редактирования активно используется буфер обмена.

Выполнять операции простейшего редактирования во всех аудиоредакторах очень просто. На рисунке 2 показано окно программы Cool Edit Pro с загруженным в нее файлом, содержащим запись стереофонического музыкального сигнала (Queen, "Богемская рапсодия").

Выделение фрагмента сигнала осуществляется точно так же, как в текстовых и графических редакторах - с помощью левой кнопки мыши. Позицию курсора указывает желтый треугольничек на линейке и вертикальная линия желтого цвета. Работать можно одновременно с двумя каналами или с каждым по отдельности - в зависимости от решаемой задачи.

К звуковым процессам обычно относят такие операции, как слияние волновых форм, инверсию, различные амплитудные преобразования, нормализацию, частотную коррекцию и преобразование тембров.

Слияние волновых форм - это, по сути, смешивание двух звуков по определенным правилам. Непродуманное слияние волновых форм всегда приводит к одному-единственному результату - жуткой какофонии, но если все сделать правильно, можно получить очень интересные эффекты.

Инверсия - это изменение фазы аудиосигнала на 180 градусов. Инвертирование монофонического сигнала никакого эффекта не дает, а вот для стерео - другое дело. Если один канал стереозаписи по каким-то причинам записан в противофазе, при воспроизведении могут пропасть басы или вообще нарушиться стереобаза. Инверсия позволяет устранить этот дефект.

К амплитудным преобразованиям относят, например, постепенное нарастание или затухание звукового сигнала, происходящее по определенному закону. Это очень простая операция.

Очень важная и полезная вещь - нормализация. Она необходима для получения максимально возможной громкости сигнала без потери его пиковых значений из-за ограничения усиления. Нормализация сигнала может выполняться по пиковым уровням или по средней мощности. Если вы собираете свой музыкальный альбом с разных компакт-дисков, виниловых пластинок или с магнитной ленты, обратите внимание на то, чтобы все композиции были нормализованы с одинаковыми параметрами, в противном случае слушать такой диск будет неприятно.

Частотная коррекция сигналов осуществляется с помощью программных эквалайзеров (рис. 3), которые очень похожи на аппаратные, имеющиеся на обычных музыкальных центрах.

К звуковым эффектам обычно относят эффекты, основанные на задержке сигнала, "Дилей", реверберацию, эхо и эффекты, построенные на модуляции сигнала с небольшой частотой (до 16 Гц). Амплитудную модуляцию еще называют амплитудным вибрато или тремоло, эффект от которой заключается в своеобразном дрожании звука.

Частотная модуляция сигнала на слух воспринимается как завывание или плавание звука. Примерно такой эффект можно получить, слегка прижав пальцем приемный вал кассетного или катушечного магнитофона.

Эффект "Вау-Вау", напоминающий открывание и закрывание некой акустической заслонки, получается благодаря тембровой модуляции сигнала.

На разных видах модуляции сигнала основаны также хорошо знакомые любителям рок-музыки эффекты "Флэнжер", "Фэйзер" и "Хорус".

К дополнительным инструментам аудиоредакторов относят анализаторы спектра, шумоподаватели и модули синхронизации аудиосигнала с видеорядом.

Представление аудиофайла в виде спектра нередко оказывается очень полезным. Преобразование звуковой волны из временной области в частотную

в аудиоредакторах осуществляется с помощью быстрого преобразования Фурье (Fast Fourier Transform, FFT). Классическая спектрограмма аудиосигнала имеет вид, показанный на рисунке 4. Здесь по оси абсцисс откладывается частота в герцах, а по оси ординат - амплитуда сигнала в децибелах.

Интересной разновидностью спектрограмм, используемых в аудиоредакторах, являются сонограммы. На сонограмме (рис. 5) по оси абсцисс откладывается время, а по вертикали - частота. Амплитуда каждой частотной составляющей представлена в условных цветах. Сонограммы бывают полезны для определения отличий спектра рабочей фонограммы от эталонных образцов спектра.

В тихих местах фонограмм и в паузах обычно бывает слышен своеобразный шипящий звук - это паразитный шум. Для борьбы с ним в аудиоредакторах используют специальные шумоподавители и разнообразные фильтры. Для того чтобы шумоподаватель мог эффективно работать, его нужно настраивать на каждую фонограмму, поскольку спектры шумов у разных записей различаются. С этой целью выбирают участок, не содержащий полезного сигнала, и по нему программа набирает статистику о характере шумов. Если на спектрограмме удалось выделить спектр шумового сигнала, то на его подавление можно настроить соответствующий фильтр. Нужно только помнить, что неумелое использование фильтров может непоправимо испортить фонограмму, поэтому эксперименты лучше проводить на заранее созданной копии.

При записи аудиофайла с виниловой пластинки нередко встречаются проблемы, связанные с характерными щелчками и потрескиваниями, возникающими при проигрывании старых дисков. При загрузке такого файла в аудиоредактор хорошо видно, чем вызваны эти щелчки (рис. 6).

Вертикальный выброс чуть правее линии курсора - это и есть графическое отображение щелчка, вызванного царапиной на пластинке. Аудиоредакторы, как правило, комплектуются специальными плагинами, способными сводить к минимуму шумы винила. Впрочем, некоторые считают, что в шорохе и потрескивании старых записей есть своеобразный шарм.

Современные аудиоредакторы - это мощные и одновременно простые и понятные, дружелюбные к пользователю программы, открывающие для любителей музыки широкое поле для творчества. Пожелаем им успехов.

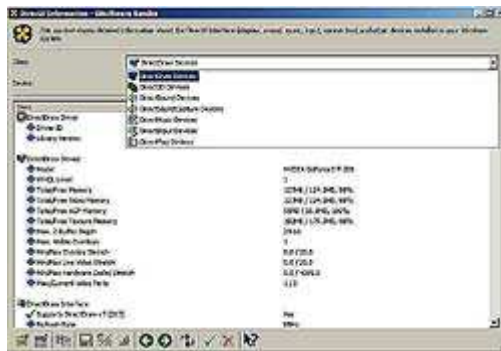


Рис. 1. Так выглядит информация о компонентах DirectX в окне программы SiSoftware Sandra



Рис. 2. Окно программы Cool Edit Pro

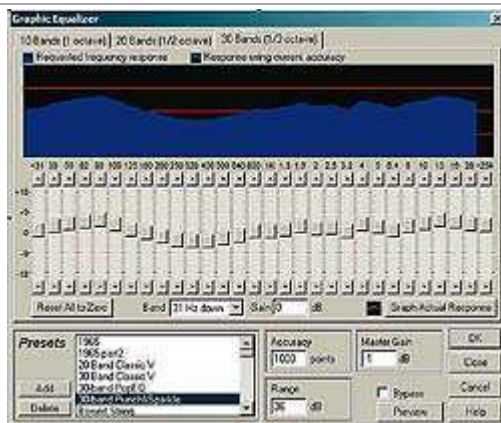


Рис. 3. Графический эквалайзер программы Cool Edit Pro

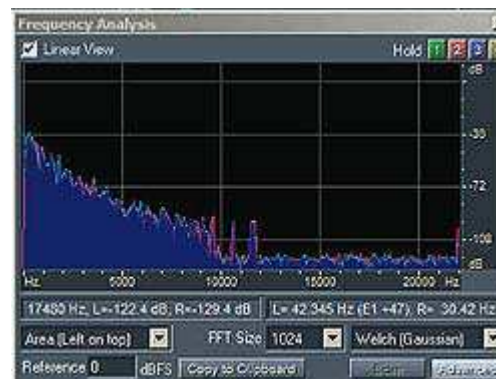


Рис. 4. Спектральное представление сигнала после FFT-преобразования



Рис. 5. Сонограмма стереоаудиофайла



Рис. 6. Аудиофайл, записанный с виниловой пластинки

