

Звуковая информация

**Учитель
информатики
Лашина Т.С.**

1. Что такое звук?

Компьютер широко применяют в настоящее время в различных сферах. Не стала исключением и обработка звуковой информации, музыка.

До 1983 года все записи музыки выходили на виниловых пластинках и компакт-кассетах. В настоящее время широкое распространение получили компакт-диски.

С начала 90-х годов ПК получили возможность работать средствами мультимедиа, в том числе со звуковой информацией.

Современная работа на ПК немыслима сегодня без средств мультимедиа. Каждый ПК, имеющий звуковую плату, микрофон и звуковые колонки, может записывать, сохранять и воспроизводить звуковую информацию.

Существуют спец. программные средства для работы со звуком, открывающие широкие возможности по созданию, редактированию и прослушиванию звуковых файлов. Созданы программы распознавания речи, и в результате появляется возможность управления ПК при помощи голоса. Недалек тот день, когда, наряду с клавиатурой, у пользователя появится возможность вводить текстовую информацию с микрофона голосом.

Каким же образом компьютер работает со звуковой информацией?

Звуковые сигналы в окружающем нас мире необычайно разнообразны. Мир наполнен самыми разнообразными звуками: тиканье часов и гул моторов, завывание ветра и шелест листьев, пение птиц и голоса людей.

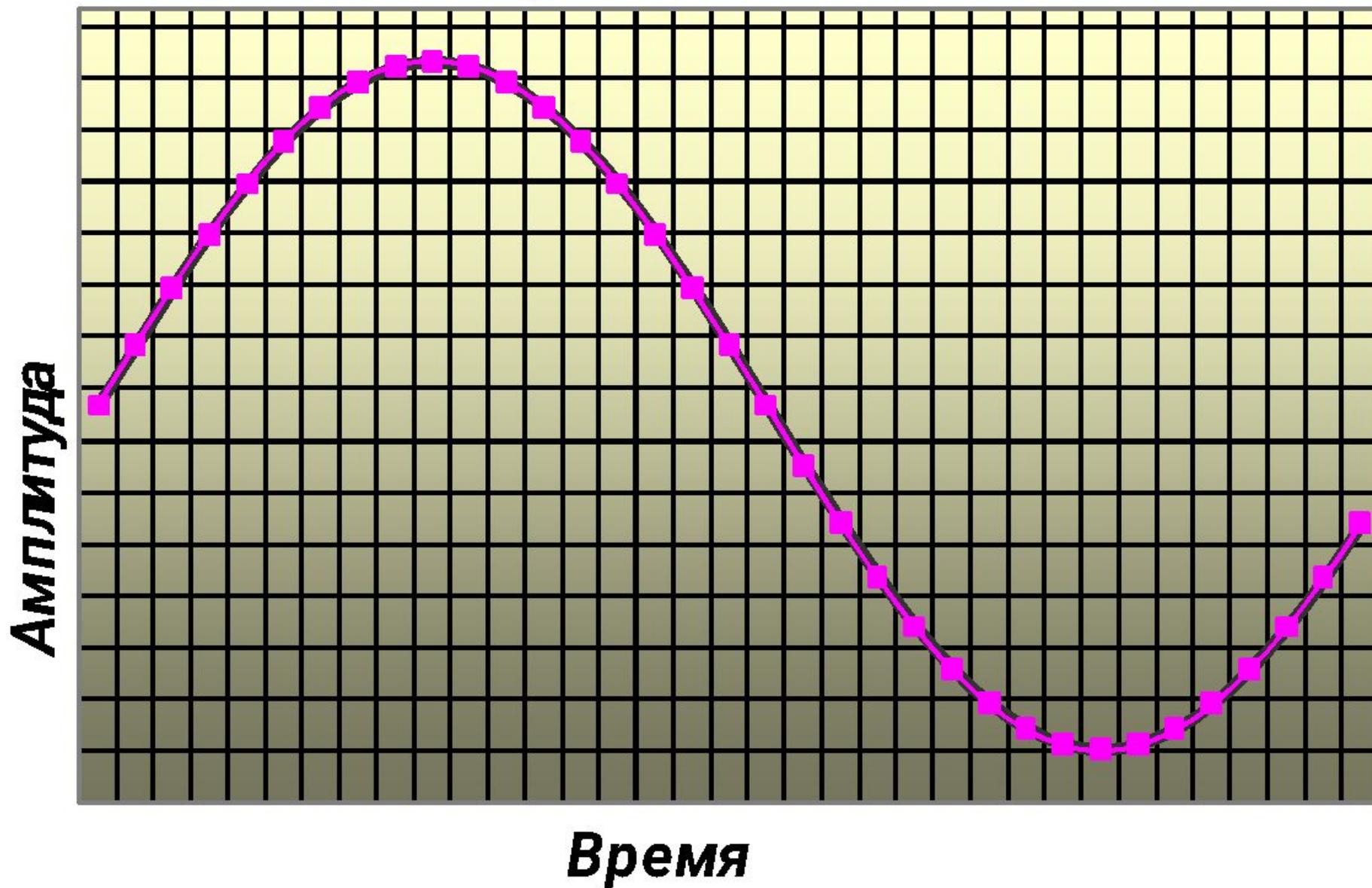
Физическая природа звука – колебания в определенном диапазоне частот, передаваемые звуковой волной через воздух.

На слух человек воспринимает упругие волны, имеющие частоту где-то в пределах от 16 Гц до 20 кГц (1 Гц - 1 колебание в секунду).

В соответствии с этим упругие волны в любой среде, частоты которых лежат в указанных пределах, называют звуковыми волнами или просто звуком.

Для их записи с целью последующего воспроизведения необходимо как можно точнее сохранить форму кривой зависимости интенсивности звука от времени.

Фрагмент звукового сигнала



При этом возникает одна очень важная и принципиальная трудность: всякий реальный звук, будь то игра музыкальных инструментов или голос человека, - это своеобразная смесь многих гармонических колебаний воздуха с определенным набором частот, т.е. звук является аналоговым сигналом, так как амплитуда этих колебаний непрерывно меняется со временем.

Компьютер же способен хранить в памяти и осуществлять обработку информации, представленную в цифровом двоичном виде, т.е. имеет дело с дискретными сигналами, изменяющимися скачкообразно («лесенкой») и принимающими конечное множество значений. Отсюда следует, что в процессе сохранения звуковой информации она должна быть "оцифрована", т.е. из аналоговой непрерывной формы переведена в цифровую дискретную.

Процесс преобразования из аналоговой формы в цифровую называют дискретизацией или оцифровкой.

2. Преобразование звуковых сигналов

Для преобразования аналоговых сигналов внешнего мира для компьютеров разработано много устройств.

Аналоговые данные могут быть преобразованы в цифровую форму и наоборот специальными устройствами, называемыми аналого-цифровыми преобразователями (АЦП) и цифро-аналоговыми преобразователями (ЦАП).

Функции этих устройств в компьютере в частности реализованы в специальном устройстве, подключаемом к компьютеру и называемым аудиоадаптером (или звуковой платой или саундбластером).

Давайте посмотрим, какой путь придется преодолеть звуковому сигналу, записываемому через микрофон в память ПК и воспроизводимому затем через колонки:



Так каков же принцип кодирования звуковой информации, какие функции выполняет аудиоадаптер?

В процессе кодирования фонограммы через равные промежутки времени, называемые семплами, аудиоадаптером производится измерение амплитуды напряжения, выходящего с микрофона, которое в свою очередь зависит от интенсивности звуковой волны. Каждой измеренной величине амплитуды присваивают значение уровня громкости звука из заданного набора уровней громкости. Это следует понимать так, что при измерении имеется «сетка» стандартных уровней.

Количество уровней громкости, т.е. число горизонтальных линий сетки зависит от разрядности аудиоадаптера. Чем выше разрядность, тем большее количество уровней громкости будет содержаться в наборе и соответственно будет более точным сохранение амплитуды звуковой волны и соответственно более качественным звучание.

Первая характеристика качества преобразования звука - число бит (разрядов), используемых для кодировки уровня громкости, называемая глубиной кодирования звука.

Если аудиоадаптер 8-ми разрядный, то с помощью него можно сохранять 2 в 8 -ой степени $=256$ уровней громкости. Это значит, что любой измеренной амплитуде будет присвоено одно из 256 значений. Очевидно, что 16 -ти разрядный (65536 уровней громкости) аудиоадаптер точнее кодирует и воспроизводит звук, чем 8 -и разрядный.

Вторая характеристика качества преобразования звука - промежутки времени между двумя измерениями амплитуды звуковой волны. Этот временной интервал называется семплом, а количество измерений входного сигнала за 1 секунду называется частотой дискретизации и измеряется в Гц.

1 измерение за 1 сек соответствует частоте 1 Гц. 1000 измерений за 1 секунду – 1 кГц.

Значения уровней громкости для каждого измерения, перечисленные по порядку и представленные n-разрядными двоичными кодами будут кодировать поступивший в память ЭВМ звуковой сигнал (n-разрядность аудиоадаптера).

Мы рассмотрели с вами универсальный способ кодирования звуковой информации, хотя существуют и другие.

Качество такой оцифровки звука определяется, следовательно 2-мя характеристиками аудиоадаптера: разрядностью (глубиной кодирования звука) и частотой дискретизации.

Именно от разрядности зависит достоверность восстановления формы волны, а следовательно окраска (тембр) звука.

Разрядность указывает с какой точностью происходят изменения амплитуды аналогового сигнала. Точность, с которой при оцифровке передается значение амплитуды сигнала в каждый из моментов времени, определяет качество сигнала после цифро-аналогового преобразования.

Именно от разрядности зависит достоверность восстановления формы волны.

Вопрос о выборе частоты дискретизации далеко не праздный, т. к. от нее зависит такое качество звука, как высота.

Человеческое ухо способно различать частоты от 30 Гц до 20 кГц. Если частота не будет более чем в два раза превышать частоту верхней границы слышимого звукового диапазона, то на высоких частотах будут происходить потери.

Количество измерений сигнала в секунду должно быть больше, чем количество колебаний за тот же промежуток времени. Если же частота дискретизации будет значительно ниже частоты звуковой волны, то амплитуда сигнала успеет несколько раз измениться за время между измерениями, а это приведет к тому, что цифровой отпечаток будет нести хаотичный набор данных. При цифро-аналоговом преобразовании такой семпл не передаст основной сигнал, а только выдаст шум.

Существует специальная теорема Найквиста, согласно которой частота оцифровки звука должна как минимум в 2 раза превышать максимальную частоту, входящую в состав спектра сигнала.

Частота может принимать значения от 8 до 48 кГц. Например, при записи речи вполне достаточно частоты дискретизации 8 кГц. Результат при этом получается хотя и не блестящий, но речь будет легко разборчива, примерно такая, как по телефону.

При записи на компакт-диски используются глубина кодирования, равная 16 разрядам и частота дискретизации, равная 44032 Гц. При частоте в 8 кГц – качество звука соответствует качеству радиотрансляции; при частоте 11 кГц – как у телефона; при частоте 22 кГц – как у кассетного магнитофона; при частоте 44 кГц – как у компакт-диска.

На современные цифровые звуковые устройства устанавливаются 20-битные преобразователи. Звук так и остается 16-битным, преобразователи повышенной разрядности устанавливают для улучшения качества записи на низких уровнях. Таким образом, чем больше разрядность и частота дискретизации, тем точнее представляется звук в цифровой форме, но тем и большие ресурсы компьютера затрачиваются на это представление.

Весьма высокие еще для начала 90-х параметры цифрового звука "16 бит/44.1 кГц" сейчас могут считаться лишь минимально допустимыми для понятий "качественный звук" и "Hi-Fi".

В студийной работе происходит переход на стандарт "24 бита/96 кГц", который по теоретически достижимому качеству пока заметно перекрывает возможности существующих звуковых систем.

Внутри стандарта "компакт-диск", ограниченного своими 16 разрядами и 44.1 кГц частоты дискретизации, используется преобразование цифрового звука под большую частоту дискретизации и разрядность с последующей интерполяцией промежуточных значений.

Само по себе это не улучшает качества звука, однако позволяет заметно снизить погрешности, возникающие из-за неидеальности ЦАП, фильтров и прочих элементов тракта.

3. Задачи.

1) УСЛОВИЕ: определите количество уровней громкости, если разрядность звуковой платы 5 бит

РЕШЕНИЕ: $2^5 = 32$

ОТВЕТ: 32 уровня громкости

2) УСЛОВИЕ: определите объем памяти для хранения цифрового аудиофайла, время звучания которого составляет 1 минута при частоте дискретизации 44032 Гц и разрядности 8 бит (ответ дайте в Кб)

3) УСЛОВИЕ: определить длительность звучания цифрового аудиофайла, занимающего объем 0,5 Мб, если он оцифрован с частотой 44 кГц 16-и разрядным аудиоадаптером (ответ дать в сек, округлив до целого значения)

4) УСЛОВИЕ: определите ширину потока звукового файла, если заданный набор громкостей составляет 128 уровней, а частота дискретизации 8 кГц

(ответ дать в бит/сек)

ОТВЕТЫ:

2) РЕШЕНИЕ: ширина потока = $44032 \text{ Гц} \times 8 \text{ бит} = 352256 \text{ бит/сек}$

объем файла = $352256 \text{ бит/сек} \times 60 \text{ сек} = 21135360 \text{ бит} = 2580 \text{ Кб}$

ОТВЕТ: 2580 кБ

3) РЕШЕНИЕ: ширина потока = $44000 \text{ Гц} \times 16 \text{ бит} = 704000 \text{ бит/сек}$

длительность = $0,5 \times 2^{23} / 704000 = 5,95 \text{ сек} \sim 6 \text{ сек}$

ОТВЕТ: 6 сек

4) РЕШЕНИЕ: так как уровней – 128, то разрядность аудиоадаптера – 7 бит ($2^7 = 128$)

ширина потока = $8000 \text{ Гц} \times 7 \text{ бит} = 56000 \text{ бит/сек}$

ОТВЕТ: 56000 бит/сек

4. Основные виды звуковых компьютерных файлов.

Оцифрованный звук может быть сохранен во внешней памяти в виде звукового файла.

Звуковой файл – это файл, хранящий звуковую информацию в числовой двоичной форме. Звуковые файлы в свою очередь тоже бывают разных типов в зависимости от способа кодирования, метода извлечения звуков.

Один из самых распространенных типов звуковых файлов – это waveAudio, или коротко wav . Данный формат обеспечивает очень хорошее качество, но имеет большой объем звукового файла. 74 минуты звучания такого формата занимают 640 Мбайт, то есть объем лазерного диска.

Формат waveAudio используется в музыкальных лазерных дисках. Чтобы уменьшить объем, занимаемый оцифрованным звуковым сигналом, применяют различные алгоритмы сжатия.

В последние годы широкое распространение получил формат MP3, обеспечивающий несколько меньшее качество по сравнению с форматом .wav , но и объем такой формат занимает примерно в 10 раз меньше. Это позволяет на одном лазерном диске уместить музыку общей длительности не 74 минуты, а целых 740 минут!

Оба вышеперечисленных типа файлов могут содержать любую звуковую информацию - музыку, речь, природные и другие звуки (характеризует глубину кодирования), и текущий уровень измеряемого сигнала округляется до ближайшего из них.

Изложенный метод преобразования звуковой информации для хранения в памяти компьютера в очередной раз подтверждает уже неоднократно обсуждавшийся ранее тезис: любая информация для хранения в компьютере приводится к цифровой форме и затем переводится в двоичную систему.

Теперь мы знаем, что и звуковая информация не является исключением из этого фундаментального правила.

характеристики. Виды программ.

Колебание, которое имеет наиболее низкую частоту, называют основным тоном, другие - обертонами.

Тембр - разное количество обертонов, присущих тому или иному звуку, которое придает ему особую окраску.

Отличие одного тембра от другого обусловлено не только числом, но и интенсивностью обертонов, сопровождающих звучание основного тона. Именно по тембру мы легко можем отличить звуки рояля и скрипки, гитары и флейты, узнать голос знакомого человека.

Музыкальный звук можно характеризовать тремя качествами: тембром, т. е. окраской звука, которая зависит от формы колебаний, высотой, определяющейся числом колебаний в секунду (частотой), и громкостью, зависящей от интенсивности колебаний.

Если имеется компьютер, на котором установлена студийная звуковая плата, с подключенными к ней MIDI-клавиатурой и микрофоном, то можно работать со специализированным музыкальным программным обеспечением.

Условно его можно разбить на несколько видов:

- 1) всевозможные служебные программы и драйверы, предназначенные для работы с конкретными звуковыми платами и внешними устройствами;
- 2) аудиоредакторы, которые предназначены для работы со звуковыми файлами, позволяют производить с ними любые операции - от разбиения на части до внесения спецэффектов;
- 3) программные синтезаторы, которые появились сравнительно недавно и корректно работают только на мощных компьютерах. Они позволяют экспериментировать с созданием различных звуков; и другие.

К первой группе относятся все служебные программы операционной системы.

Так, например, win 95 и 98 имеют свои собственные программы микшеры и утилиты для воспроизведения или записи звука, проигрывания компакт-дисков и стандартных MIDI - файлов. Установив звуковую плату можно при помощи этих программ проверить ее работоспособность.

Например, программа Фонограф предназначена для работы с wave-файлами (файлы звукозаписи в формате Windows). Эти файлы имеют расширение .WAV .

Эта программа предоставляет возможность воспроизводить, записывать и редактировать звукозапись приемами, аналогичными приемам работы с магнитофоном. Желательно для работы с Фонографом подключить микрофон к компьютеру. Если необходимо сделать звукозапись, то нужно определиться с качеством звука, так как именно от нее зависит продолжительность звукозаписи. Возможная продолжительность звучания тем меньше, чем выше качество записи. При среднем качестве записи можно удовлетворительно записывать речь, создавая файлы продолжительностью звучания до 60 секунд. Примерно 6 секунд будет продолжительность записи, имеющая качество музыкального компакт - диска.

Но для того чтобы записать звук на какой-нибудь носитель его нужно преобразовать в электрический сигнал. Это делается с помощью микрофона.

Самые простые микрофоны имеют мембрану, которая колеблется под воздействием звуковых волн. К мембране присоединена катушка, перемещающаяся синхронно с мембраной в магнитном поле. В катушке возникает переменный электрический ток. Изменения напряжения тока точно отражают звуковые волны.

Звуковые волны при помощи микрофона превращаются в аналоговый переменный электрический сигнал.

Переменный электрический ток, который появляется на выходе микрофона, называется аналоговым сигналом.

Он проходит через звуковой тракт и попадает в аналого-цифровой преобразователь (АЦП) - устройство, которое переводит сигнал в цифровую форму. Если использовать 8-битное кодирование, то можно достичь точность изменения амплитуды аналогового сигнала до $1/256$ от динамического диапазона цифрового устройства ($2^8 = 256$).

Если использовать 16-битное кодирование для представления значений амплитуды звукового сигнала, то точность измерения возрастет в 256 раз. В современных преобразователях принято использовать 20-битное кодирование сигнала, что позволяет получать высококачественную оцифровку звука.

Благодаря высокому качеству цифровая техника стремительно «врывается» в наш дом. Согласно прогнозам, в 21 веке большая часть бытовой аудио- и видеотехники будет цифровой.

6. Задачи.

1. Подсчитать, сколько места будет занимать одна минута цифрового звука на жестком диске или любом другом цифровом носителе, записанного с частотой

а) 44.1 кГц; б) 11 кГц; в) 22 кГц; г) 32 кГц

и разрядностью 16 бит.

Решение:

а) Если записывают моносигнал с частотой 44.1 кГц, разрядностью 16 бит (2 байта), то каждую минуту аналого-цифровой преобразователь будет выдавать $441000 * 2 * 60 = 529000$ байт (примерно 5 Мб) данных об амплитуде аналогового сигнала, который в компьютере записывается на жесткий диск.

Если записывают стереосигнал, то 1058000 байт (около 10 Мб)

б) для частот 11, 22, 32 кГц расчеты производятся аналогично.

2. Какой информационный объем имеет моноаудиофайл, длительность звучания которого 1 секунда, при среднем качестве звука (16 бит, 24 кГц)?

Решение:

$$16 \text{ бит} * 24000 = 384000 \text{ бит} = 48000 \text{ байт} = 47 \text{ кБайт}$$

3. Рассчитайте объем стереоаудиофайла длительностью 20 секунд при 20-битном кодировании и частоте дискретизации 44.1 кГц.

Решение:

$$20 \text{ бит} * 20 * 44100 * 2 = 35280000 \text{ бит} = 4410000 \text{ байт} = 4.41 \text{ Мб}$$

4. Определить количество уровней звукового сигнала при использовании устаревших 8-битных звуковых карт.

Решение: $K = 2^8 = 256.$