

Задание: изучить теоретический материал и составить конспект.
Представление звуковой информации в компьютере

Звук представляет собой непрерывный сигнал — звуковую волну с меняющейся амплитудой и частотой. Чем больше амплитуда сигнала, тем он громче для человека. Чем больше частота сигнала, тем выше тон. Частота звуковой волны выражается числом колебаний в секунду и измеряется в герцах (Гц, Hz). Человеческое ухо способно воспринимать звуки в диапазоне от 20 Гц до 20 кГц, который называют звуковым.

Количество бит, отводимое на один звуковой сигнал, называют глубиной кодирования звука. Современные звуковые карты обеспечивают 16-, 32- или 64-битную глубину кодирования звука.

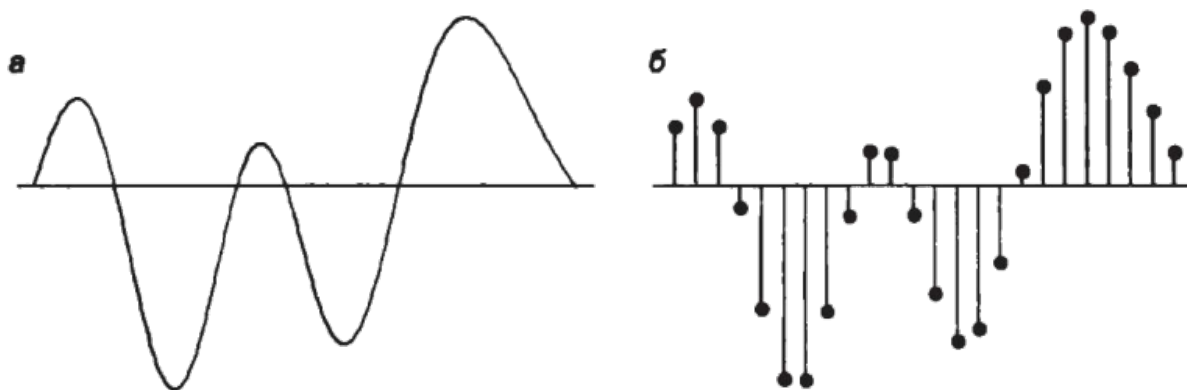
При кодировании звуковой информации непрерывный сигнал заменяется дискретным, то есть превращается в последовательность электрических импульсов (двоичных нулей и единиц). Важной характеристикой при кодировании звука является частота дискретизации — количество измерений уровней сигнала за 1 секунду:

- 1 (одно) измерение в секунду соответствует частоте 1 Гц;
- 1000 измерений в секунду соответствует частоте 1 кГц.

Количество измерений может лежать в диапазоне от 8 кГц до 48 кГц (от частоты радиотрансляции до частоты, соответствующей качеству звучания музыкальных носителей).

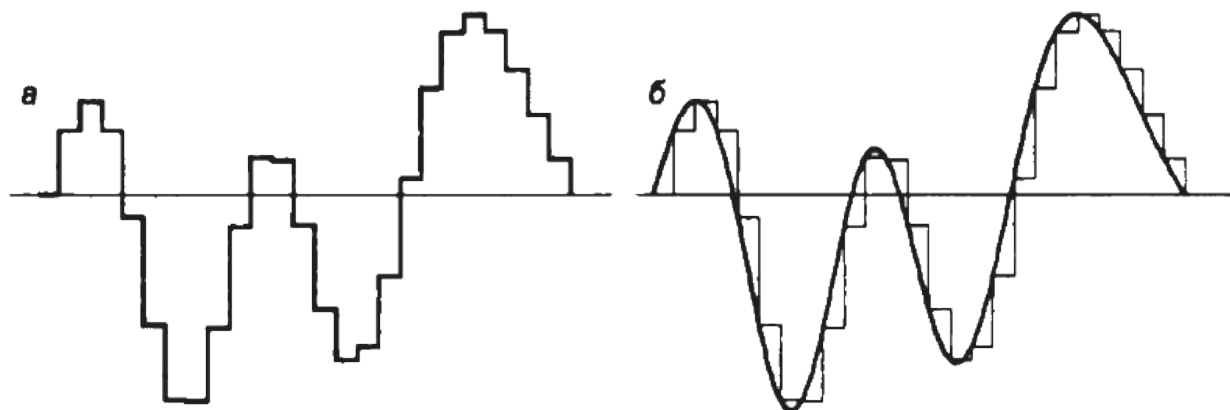
Существуют различные методы кодирования звуковой информации двоичным кодом, среди которых можно выделить два основных направления: метод FM и метод Wave-Table.

Метод FM (Frequency Modulation) основан на том, что теоретически любой сложный звук можно разложить на последовательность простейших гармонических сигналов разных частот, каждый из которых представляет собой правильную синусоиду, и следовательно, может быть описан кодом. Разложение звуковых сигналов в гармонические ряды и представление в виде дискретных цифровых сигналов (рис. 1.5) выполняют специальные устройства — аналогово-цифровые преобразователи (АЦП).



*Рис. 1. Преобразование звукового сигнала в дискретный сигнал:
а — звуковой сигнал на входе АЦП; б — дискретный сигнал на выходе АЦП*

Обратное преобразование для воспроизведения звука, закодированного числовым кодом, выполняют цифро-аналоговые преобразователи (ЦАП). Процесс преобразования звука представлен на рис. 2. Данный метод кодирования не дает хорошего качества звучания, но обеспечивает компактный код.



*Рис 2. Преобразование дискретного сигнала в звуковой сигнал:
а — дискретный сигнал на входе ЦАП; б — звуковой сигнал на выходе ЦАП*

Таблично-волновой метод (Wave-Table) основан на том, что в заранее подготовленных таблицах хранятся образцы звуков окружающего мира, музыкальных инструментов и т. д. Числовые коды выражают высоту тона, продолжительность и интенсивность звука и прочие параметры, характеризующие особенности звука. Поскольку в качестве образцов используются «реальные» звуки, качество звука, полученного в результате синтеза, получается очень высоким и приближается к качеству звучания реальных музыкальных инструментов.

Звуковые файлы имеют несколько форматов. Наиболее популярные из них MIDI, WAV, MP3.

Формат MIDI (Musical Instrument Digital Interface) изначально был предназначен для управления музыкальными инструментами. В настоящее время используется в области электронных музыкальных инструментов и компьютерных модулей синтеза.

Формат аудиофайла WAV (waveform) представляет произвольный звук в виде цифрового представления исходного звукового колебания или звуковой волны. Все стандартные звуки Windows имеют расширение WAV.

Формат MP3 (MPEG-1 Audio Layer 3) — один из цифровых форматов хранения звуковой информации. Он обеспечивает более высокое качество кодирования.

Кодирование аудиоинформации

Так как компьютер работает с числами, звуки и музыка должны быть представлены в числовом виде, или, как принято говорить, закодированы. Произвольная аудиоинформация при кодировании занимает много места,

поэтому часто используют сжатые аудиоформаты. Музыка занимает меньше места, так как хорошо формализуется – ее можно записать с помощью нот.

Звук представляет собой волну, распространяющуюся в атмосфере, и воспринимаемую человеком с помощью органов слуха. Громкость звука – это его кажущаяся сила. Измеряется громкость в децибелах (дБ). Громкость обычного разговора около 50 дБ, шум на улице часто превышает 70 дБ, а громкость взлетающего самолета составляет 120 дБ. Порог чувствительности человеческого уха около 20 дБ.

Характеризуется звуковая волна изменением во времени частоты и амплитуды сигнала. Графически звуковая волна описывается кривой, задающей зависимость амплитуды от времени. Частота основных колебаний определяет высоту звука. Но звуки одной частоты могут иметь разный тембр.

Чтобы закодировать звук, необходимо измерять амплитуду сигнала через определенные промежутки времени. На каждом временном отрезке определяется средняя амплитуда сигнала. Графически такое преобразование описывается множеством столбиков.

При восстановлении исходной кривой ее вид будет искажен. Искажения тем больше, чем больше ширина столбиков, то есть чем реже определяется текущая амплитуда. Чем промежутки времени меньше, тем выше будет качество закодированного звука. Частота, с которой определяется амплитуда сигнала, называется частотой дискретизации.

Амплитуда сигнала, определенная в каждый момент времени, также должна быть представлена в числовом виде. В простейшем случае можно использовать один бит – есть звук или его нет. Но на практике такое кодирование не имеет смысла. Минимально для кодирования амплитуды сигнала отводятся восемь бит – один байт, что позволяет описать двести пятьдесят шесть уровней громкости. Качество звука при этом получается не слишком высокое. Если и частота дискретизации невелика, то при воспроизведении будут присутствовать сильные искажения. Значительно лучшее качество получается при использовании двух байт, что позволяет задать более шестидесяти пяти тысяч разных значений амплитуды. В большинстве случаев двух байт достаточно для получения высококачественной записи звука, хотя иногда применяют 24 бита – три байта для кодирования амплитуды сигнала.

Для кодирования звуков следует использовать частоту вдвое большую, чем частота кодируемого звука. Объяснение этому довольно простое. Звуковая волна состоит из двух полупериодов: положительного и отрицательного. Поэтому для ее имитации необходимо иметь хотя бы по одной выборке на каждом из полупериодов. Так как человек воспринимает звуки в диапазоне частот от 20 до 20000 Гц, то для качественного кодирования необходимо использовать частоту вдвое большую, чем 20000, то есть 40000 Гц. Тогда сохраненные выборки позволят воспроизводить звуковую волну внутри диапазона, воспринимаемого человеческим ухом. Для качественного кодирования звука принято иметь некоторый запас, поэтому при цифровой звукозаписи используется частота дискретизации 44100 Гц и 48000 Гц. Это

означает, что за каждую секунду звукозаписи в цифровом виде записывается более 44000 единиц информации, последовательность которых моделирует звук длительностью в одну секунду.

Для того, чтобы записать стереозвук, следует одновременно кодировать два независимых канала звука. При этом, чтобы получить хорошее качество, нужно использовать два байта для кодирования и частоту дискретизации 44100 Гц для каждого из каналов. Именно так кодируется звук на компакт-дисках. При этом одна минута закодированного звука займет более 10 Мб. В некоторых случаях можно обойтись более низким качеством, сравнимым с качеством записи диктофона. Для того чтобы закодировать голос, не предъявляя повышенных требований к качеству звучания, можно использовать один байт при кодировании и один монофонический канал. Частоту дискретизации также можно понизить. Чтобы различать отдельные слова и понимать их смысл, достаточно частоты дискретизации 8000 Гц. С такими параметрами минута закодированного звука займет менее 480 Кб.

Для повышения качества кодирования используют более высокие частоты дискретизации, до 96000 Гц, однако такое качество требуется исключительно при работе в профессиональных звукозаписывающих студиях.

Современные компьютеры часто используются при создании и воспроизведении музыки. Музыкальное произведение можно закодировать как любой другой звук, однако это займет много места. Кроме того, возникнут трудности при изменении партий отдельных инструментов. Проще указать инструмент и задать, какую ноту и как долго он должен играть. Для воспроизведения музыки компьютер синтезирует разнообразные звуки, которые издают музыкальные инструменты.

В компьютерной музыке используется аббревиатура MIDI, которая расшифровывается как Musical Instrument Digital Interface (Цифровой интерфейс музыкальных инструментов). Имеется стандарт, описывающий основные используемые инструменты, – GM (General MIDI – единый MIDI). В стандарте описаны пятнадцать групп мелодических инструментов и одна группа ударных инструментов. Мелодический набор состоит из пианино, органов, гитар, струнных, духовых и тому подобных инструментов. За всеми инструментами закреплены номера, например, нулевой номер имеет акустический рояль. Кроме GM используются стандарты GS (General Synth – единый синтез), XG (Extended General – единый расширенный), GM2 (General MIDI 2). Все эти стандарты не заменяют собой GM, а лишь дополняют его новыми инструментами и дополнительными параметрами звучания.

Несмотря на то, что инструменты и тембры стандартизированы в GM, а MIDI-файл содержит только номера инструментов и тембров, этот файл по-разному будет воспроизводиться на разных звуковых картах. Это объясняется несколькими причинами. Так, в стандарте описаны только названия инструментов и тембров. Такие параметры звука, как громкость, окраска и другие не определены и выбираются производителями звуковых карт произвольно.

Кроме того, на качество воспроизведения звука сильно сказывается метод, которым этот звук воспроизводится. Применяют два основных метода синтеза звуков. Более простой метод называется частотным синтезом (FM-синтез). Для каждой ноты каждого инструмента определена частота и амплитуда звука, и звуковая плата компьютера синтезирует звук. Однако при этом синтезированные звуки получаются не слишком похожими на звучание реальных инструментов. В современных звуковых платах частотный синтез не используется.

Значительно лучшее качество звучания дают волновые таблицы (Wave Table). В таблице записаны закодированные звуки реальных инструментов. При этом используется метод кодирования амплитуды звукового сигнала через короткие промежутки времени. Например, если требуется воспроизвести удар по тарелке, звуковая плата проигрывает небольшой фрагмент, записанный в определенном месте таблицы. Фрагменты называют сэмплами (samples). Инструменты с малой длительностью звучания обычно записываются полностью, а для остальных может записываться лишь начало, конец звука и небольшая средняя часть, которая затем проигрывается в цикле в течение нужного времени. Такое кодирование обеспечивает предельную реалистичность звучания классических инструментов и простоту получения звука. Однако волновые таблицы могут занимать много места в памяти.

Так как музыка, представленная в цифровом виде, не требует преобразований, к компьютеру напрямую можно подключить цифровые синтезаторы. Наигрывая мелодию на синтезаторе, в компьютер вводится последовательность нот. Также синтезаторы позволяют проигрывать композиции, созданные на компьютере. Загрузив в синтезатор сэмплы из волновой таблицы, можно извлекать самые необычные звуки при нажатии клавиш.

В последнее время стало модным караоке, и в компьютере стали кодировать музыку вместе с текстом. Фактически караоке является вариантом MIDI. Музыка закодирована обычным способом, но дополнительно добавлен текст, заменивший описание одного из инструментов.

Хотя частота дискретизации при кодировании звукового сигнала по компьютерным меркам не очень велика, объем получившихся цифровых данных достаточно большой. Чтобы уменьшить объем, занимаемый цифровыми аудиоданными, применяют различные методы сжатия информации, в частности алгоритмы MPEG. Например, применение сжатия по алгоритму MPEG-1 Layer 3 (MP3) позволяет уменьшить объем данных более чем в десять раз, при сохранении качества звука, близкого к audio-CD. Наряду с MP3 применяется формат сжатия по стандарту WMA (Windows Media Audio), поддерживаемый последними версиями операционных систем Windows.

В обоих стандартах используется метод сжатия по психоакустической модели, то есть из исходного звукового сигнала удаляется информация, малозаметная на слух, после чего сигнал сжимается обычными методами, которые реализованы в программах-архиваторах. При таком методе

кодирования неизбежно искажение исходного сигнала, а значит – потеря качества. Степень потери качества можно регулировать, однако при увеличении качества неизбежно растет объем информации. Основным параметром, характеризующим качество записи, является скорость потока данных, поступающих для декодирования. Часто этот параметр называют битрейтом (bitrate – частота битов).

Битрейт измеряется в килобитах в секунду и может составлять до 320 Кбит/с. В большинстве случаев вполне хватает 192 или даже 128 битрейт. Битрейт ниже 48 Кбит/с существенно ухудшит качество и его не следует применять для записи музыки. Для записи речи можно использовать меньший битрейт. Качественную диктофонную запись можно получить при битрейте равном 8 Кбит/с. Искажения при кодировании в форматах MP3 и WMA во многом зависят от характера музыки. Симфоническая музыка требует большего битрейта, а танцевальная – меньшего. Наиболее популярным битрейтом при кодировании музыкальных композиций считается битрейт 128 Кбит/с, дающий хорошее качество записи и позволяющий сжимать исходную информацию более чем в десять раз. Для хранения произвольных звуковых данных чаще всего используются файлы формата wav. В этом формате может храниться моно- или стереозвук, закодированный одним или двумя байтами и с различной частотой дискретизации. Файлы этого формата могут быть сжаты разными способами для достижения меньшего размера, а могут оставаться и несжатыми. Музыкальные файлы используют формат mid, так как цифровой музыкальный интерфейс и способ кодирования музыкальной информации называется MIDI. Сжатые файлы могут иметь расширение wav, а могут расширением указывать на используемый способ сжатия – mp3 или wma. Есть и несколько других форматов звуковых файлов, но они применяются значительно реже.