

Кодирование звуковой информации

Звук представляет собой упругие волны, распространяющиеся в какой-либо упругой среде и создающие в ней механические колебания.

В основе кодирования звука с использованием ПК лежит процесс преобразования колебаний воздуха в колебания электрического тока и последующая дискретизация аналогового электрического сигнала. Кодирование и воспроизведение звуковой информации осуществляется с помощью специальных программ (редактор звукозаписи). Качество воспроизведения закодированного звука зависит от частоты дискретизации и её разрешения (глубины кодирования звука - количество уровней)

Цифровой звук — это аналоговый звуковой сигнал, представленный посредством дискретных численных значений его амплитуд.

Оцифровка звука – это технология преобразования аналогового звукового сигнала в цифровой вид

Оцифровка звука включает в себя два процесса:

- процесс дискретизации (осуществление выборки) сигнала по времени
- процесс квантования по амплитуде.

Дискретизация по времени

Временная дискретизация заключается в разбиении непрерывной звуковой волны на отдельные маленькие временные участки.

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

Процесс дискретизации по времени — процесс получения значений сигнала, который преобразуется, с определенным временным шагом—шагом дискретизации. Количество замеров величины сигнала, осуществляемых в одну секунду, называют частотой дискретизации или частотой выборки. Чем меньше шаг дискретизации, тем выше частота дискретизации и тем более точное представление о сигнале нами будет получено.

Квантование амплитуды

Квантование (дискретизация по уровню) - это представление числа в виде цифрового кода конечной длины.

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

Разрядность кодирования - это число битов, используемое для хранения одного отсчёта.

Отведём для записи одного значения амплитуды сигнала в памяти компьютера N бит. Значит, с помощью одного N -битного слова можно описать 2^N разных положений. Каждый из 2^N возможных уровней называется уровнем квантования, а расстояние между двумя ближайшими уровнями квантования называется шагом квантования. Если амплитудная шкала разбита на уровни линейно, квантование называют линейным (однородным). Число N называют разрядностью квантования. Принимается, что погрешности квантования, являющиеся результатом квантования с разрядностью 16 бит, остаются для слушателя почти незаметными. Этот способ оцифровки сигнала — дискретизация сигнала во времени в совокупности с методом однородного

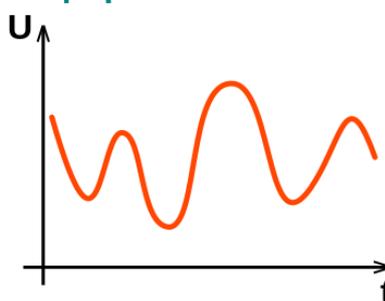
квантования — называется импульсно-кодовой модуляцией, ИКМ (англ. Pulse Code Modulation — PCM).

Оцифрованный сигнал в виде набора последовательных значений амплитуды уже можно сохранить в памяти компьютера. В случае, когда записываются абсолютные значения амплитуды, такой формат записи называется PCM (Pulse Code Modulation). Стандартный аудио компакт-диск (CD-DA), применяющийся с начала 80-х годов 20-го столетия, хранит информацию в формате PCM с частотой дискретизации 44.1 кГц и разрядностью квантования 16 бит.

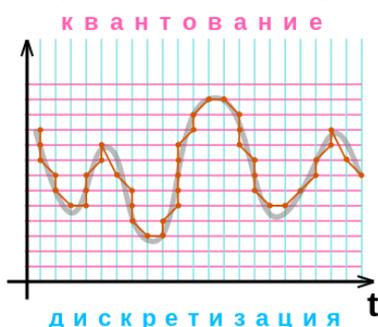
Известные форматы звуковых файлов: wav, mp3, wma, ogg, aac.

Представление аналогового сигнала в цифровой форме

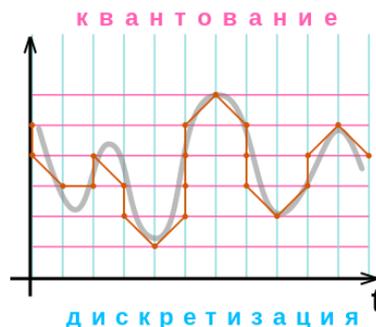
Пример аналогового сигнала



Кодирование сигнала в цифровой вид (высокое качество)



Кодирование сигнала в цифровой вид (низкое качество)



Пример задач

Производится двухканальная (стерео) звукозапись с частотой дискретизации 8 кГц и глубиной кодирования 16 бит. Запись длится 3 минуты, ее результаты записываются в файл, сжатие данных не производится. Какое из приведенных ниже чисел наиболее близко к размеру полученного файла, выраженному в мегабайтах?

Решение (через степени двойки):

1. так как частота дискретизации 8 кГц, то за одну секунду запоминается 8000 значений сигнала
2. так как глубина кодирования – 16 бит = 2 байта, то для хранения 1 секунды записи требуется 8000 x 2 байта
3. Так как производится стерео запись (записываются сразу два канала), то надо умножить на 2 8000 x 2 байта x 2
4. за 1 минуту = 60 секунд записи потребуется 8000 x 2 байта x 2 x 60
5. переходим к степеням двойки, заменяя $8 = 2^3$; $1000 \leftarrow 1024 = 2^{10}$, $60 \leftarrow 64 = 2^6$
 $2^3 \times 2^{10} \times 2 \text{ байта} \times 2 \times 2^6 = 2 \times 2^{10} \times 2^{10} = 2 \text{ Мбайта}$
6. таким образом, правильный ответ – 2.

Еще пример задачи

Проводилась одноканальная (моно) звукозапись с частотой дискретизации 16 кГц и 24-битным разрешением. В результате был получен файл размером 3 Мбайт, сжатие данных не производилось. Какая из приведенных ниже величин наиболее близка к времени, в течение которого проводилась запись?

- 1) 30 сек 2) 60 сек 3) 90 сек 4) 120 сек

Решение

- так как частота дискретизации 16 кГц, то за одну секунду запоминается 16000 значений сигнала
- так как глубина кодирования – 24 бита = 3 байта, то для хранения 1 секунды записи требуется
 $16000 \times 3 \text{ байта} = 48000 \text{ байт}$ или $= 3 \times 2^4 \times 2^{10}$
- 3 Мбайта $= 3 \times 2^{10} \times 2^{10}$ / на 1 секунду
 $3 \times 2^{10} \times 2^{10} / 3 \times 2^4 \times 2^{10} = 2^{10} / 2^4 = 2^6 = 64$ секунды
- таким образом, правильный ответ – 2.

Задачи по теме

1.	Проводилась одноканальная (моно) звукозапись с частотой дискретизации 16 кГц и 32-битным разрешением. В результате был получен файл размером 1 Мбайт, сжатие данных не производилось. Какая из приведенных ниже величин наиболее близка к времени, в течение которого проводилась запись? 1) 10 сек 2) 30 сек 3) 50 сек 4) 75 сек
2.	Проводилась одноканальная (моно) звукозапись с частотой дискретизации 16 кГц и 32-битным разрешением. В результате был получен файл размером 20 Мбайт, сжатие данных не производилось. Какая из приведенных ниже величин наиболее близка к времени, в течение которого проводилась запись? 1) 1 мин 2) 2 мин 3) 5 мин 4) 10 мин
3.	Производится четырёхканальная (квадро) звукозапись с частотой дискретизации 48 кГц и 32-битным разрешением. Запись длится 2 минуты, её результаты записываются в файл, сжатие данных не производится. Какая из приведённых ниже величин наиболее близка к размеру полученного файла? 1) 15 Мбайт 2) 27 Мбайт 3) 42 Мбайт 4) 88 Мбайт
4.	Двухканальная (стерео) звукозапись с частотой дискретизации 16 кГц и 24-битным разрешением велась в течение 5 минут. Сжатие данных не производилось. Какая из приведённых ниже величин наиболее близка к размеру полученного файла? 1) 10 Мбайт 2) 30 Мбайт 3) 50 Мбайт 4) 70 Мбайт
5.	В течение 4 минут производится двухканальная (стерео) звукозапись. Результаты записи записываются в файл, размер полученного файла - 40 Мбайт (с точностью до 10 Мбайт); сжатие данных не производилось. Среди перечисленных ниже режимов укажите тот, в котором проводилась звукозапись. 1) Частота дискретизации 16 кГц и 24-битное разрешение 2) Частота дискретизации 16 кГц и 16-битное разрешение 3) Частота дискретизации 32 кГц и 24-битное разрешение 4) Частота дискретизации 32 кГц и 16-битное разрешение