

FAQ по цифровому представлению звуковых сигналов

- [Чем цифровое представление сигналов отличается от аналогового?](#)
- [Как звук представляется в цифровом виде?](#)
- [Что такое АЦП и ЦАП?](#)
- [Что такое передискретизация \(oversampling\)?](#)
- [Как устроены и работают АЦП и ЦАП?](#)
- [Каковы достоинства и недостатки цифрового звука?](#)
- [Какие интерфейсы используются для передачи цифрового звука?](#)
- [Какие методы используются для эффективного сжатия цифрового звука?](#)
- [Какие форматы используются для представления цифрового звука?](#)
- [Как выполняется обработка цифрового звука?](#)
- [Что представляет собой цифровая звуковая рабочая станция?](#)
- [Какие жесткие диски используются в рабочих станциях?](#)
- [Есть ли преимущества у SCSI перед IDE при цифровой записи?](#)
- [Зачем в цифровой записи применяются жесткие диски класса AV?](#)
- [Искажается ли цифровой сигнал при передаче и хранении?](#)
- [Страдает ли качество сигнала при цифровой обработке?](#)
- [Страдает ли качество сигнала при преобразованиях форматов?](#)
- [Как преобразовать цифровой звук из одного формата в другой?](#)
- [Какие компьютерные программы используются для обработки звука?](#)
- [Какие платы применяются для работы с цифровым звуком на IBM PC?](#)
- [Связана ли разрядность отсчета с разрядностью канала передачи?](#)
- [Что такое джиттер?](#)
- [Может ли джиттер передаваться при копировании?](#)
- [Может ли один и тот же цифровой сигнал звучать по-разному?](#)

Автор: Евгений Музыченко (2:5000/14@FidoNet, music@spider.nrcde.ru)

Чем цифровое представление сигналов отличается от аналогового?

Традиционное аналоговое представление сигналов основано на подобии (аналогичности) электрических сигналов (изменений тока и напряжения) представленным ими исходным сигналам (звуковому давлению, температуре, скорости и т.п.), а также подобии форм электрических сигналов в различных точках усилительного или передающего тракта. Форма электрической кривой, описывающей (также говорят - переносящей) исходный сигнал, максимально приближена к форме кривой этого сигнала.

Такое представление наиболее точно, однако малейшее искажение формы несущего электрического сигнала неизбежно повлечет за собой такое же искажение формы и сигнала переносимого. В терминах теории информации, количество информации в несущем сигнале в точности равно количеству информации в исходном, и электрическое представление не содержит избыточности, которая могла бы защитить переносимый сигнал от искажений при хранении, передаче и усилении.

Цифровое представление электрических сигналов призвано внести в них избыточность, предохраняющую от воздействия паразитных помех. Для этого на несущий электрический сигнал накладываются серьезные ограничения - его амплитуда может принимать только

два предельных значения - 0 и 1. Вся зона возможных амплитуд в этом случае делится на три зоны: нижняя представляет нулевые значения, верхняя - единичные, а промежуточная является запрещенной - внутрь нее могут попадать только помехи. Таким образом, любая помеха, амплитуда которой меньше половины амплитуды несущего сигнала, не оказывает влияния на правильность передачи значений 0 и 1. Помехи с большей амплитудой также не оказывают влияния, если длительность импульса помехи ощутимо меньше длительности информационного импульса, а на входе приемника установлен фильтр импульсных помех.

Сформированный таким образом цифровой сигнал может переносить любую полезную информацию, которая закодирована в виде последовательности битов - нулей и единиц; частным случаем такой информации являются электрические и звуковые сигналы. Здесь количество информации в несущем цифровом сигнале значительно больше, нежели в кодированном исходном, так что несущий сигнал имеет определенную избыточность относительно исходного, и любые искажения формы кривой несущего сигнала, при которых еще сохраняется способность приемника правильно различать нули и единицы, не влияют на достоверность передаваемой этим сигналом информации. Однако в случае воздействия значительных помех форма сигнала может исказиться настолько, что точная передача переносимой информации становится невозможной - в ней появляются ошибки, которые при простом способе кодирования приемник не сможет не только исправить, но и обнаружить. Для еще большего повышения стойкости цифрового сигнала к помехам и искажениям применяется цифровое избыточное кодирование двух типов: проверочные (EDC - Error Detection Code, обнаруживающий ошибку код) и корректирующие (ECC - Error Correction Code, исправляющий ошибку код) коды. Цифровое кодирование состоит в простом добавлении к исходной информации дополнительных битов и/или преобразовании исходной битовой цепочки в цепочку большей длины и другой структуры. EDC позволяет просто обнаружить факт ошибки - искажение или выпадение полезной либо появление ложной цифры, однако переносимая информация в этом случае также искажается; ECC позволяет сразу же исправлять обнаруженные ошибки, сохраняя переносимую информацию неизменной. Для удобства и надежности передаваемую информацию разбивают на блоки (кадры), каждый из которых снабжается собственным набором этих кодов.

Каждый вид EDC/ECC имеет свой предел способности обнаруживать и исправлять ошибки, за которым опять начинаются необнаруженные ошибки и искажения переносимой информации. Увеличение объема EDC/ECC относительно объема исходной информации в общем случае повышает обнаруживающую и корректирующую способность этих кодов.

В качестве EDC популярен циклический избыточный код CRC (Cyclic Redundancy Check), суть которого состоит в сложном перемешивании исходной информации в блоке и формированию коротких двоичных слов, разряды которых находятся в сильной перекрестной зависимости от каждого бита блока. Изменение даже одного бита в блоке вызывает значительное изменение вычисленного по нему CRC, и вероятность такого искажения битов, при котором CRC не изменится, исчезающе мала даже при коротких (единицы процентов от длины блока) словах CRC. В качестве ECC используются коды Хэмминга (Hamming) и Рида-Соломона (Reed-Solomon), которые также включают в себя и функции EDC.

Информационная избыточность несущего цифрового сигнала приводит к значительному (на порядок и более) расширению полосы частот, требуемой для его успешной передачи, по сравнению с передачей исходного сигнала в аналоговой форме. Кроме собственно

информационной избыточности, к расширению полосы приводит необходимость сохранения достаточно крутых фронтов цифровых импульсов.

Кроме целей помехозащиты, информация в цифровом сигнале может быть подвергнута также линейному или канальному кодированию, задача которого - оптимизировать электрические параметры сигнала (полосу частот, постоянную составляющую, минимальное и максимальное количество нулевых/единичных импульсов в серии и т.п.) под характеристики реального канала передачи или записи сигнала.

Полученный несущий сигнал, в свою очередь, также является обычным электрическим сигналом, и к нему применимы любые операции с такими сигналами - передача по кабелю, усиление, фильтрование, модуляция, запись на магнитный, оптический или другой носитель и т.п. Единственным ограничением является сохранение информационного содержимого - так, чтобы при последующем анализе можно было однозначно выделить и декодировать переносимую информацию, а из нее - исходный сигнал.

Как звук представляется в цифровом виде?

Исходная форма звукового сигнала - непрерывное изменение амплитуды во времени - представляется в цифровой форме с помощью "перекрестной дискретизации" - по времени и по уровню.

Согласно теореме Котельникова, любой непрерывный процесс с ограниченным спектром может быть полностью описан дискретной последовательностью его мгновенных значений, следующих с частотой, как минимум вдвое превышающей частоту наивысшей гармоники процесса; частота F_d выборки мгновенных значений (отсчетов) называется частотой дискретизации.

Из теоремы следует, что сигнал с частотой F_a может быть успешно дискретизирован по времени на частоте $2F_a$ только в том случае, если он является чистой синусоидой, ибо любое отклонение от синусоидальной формы приводит к выходу спектра за пределы частоты F_a . Таким образом, для временной дискретизации произвольного звукового сигнала (обычно имеющего, как известно, плавно спадающий спектр), необходим либо выбор частоты дискретизации с запасом, либо принудительное ограничение спектра входного сигнала ниже половины частоты дискретизации.

Одновременно с временной дискретизацией выполняется амплитудная - измерение мгновенных значений амплитуды и их представление в виде числовых величин с определенной точностью. Точность измерения (двоичная разрядность N получаемого дискретного значения) определяет соотношение сигнал/шум и динамический диапазон сигнала (теоретически это - взаимно-обратные величины, однако любой реальный тракт имеет также и собственный уровень шумов и помех).

Полученный поток чисел (серий двоичных цифр), описывающий звуковой сигнал, называют импульсно-кодовой модуляцией или ИКМ (Pulse Code Modulation, PCM), так как каждый импульс дискретизованного по времени сигнала представляется собственным цифровым кодом.

Чаще всего применяют линейное квантование, когда числовое значение отсчета пропорционально амплитуде сигнала. Из-за логарифмической природы слуха более целесообразным было бы логарифмическое квантование, когда числовое значение

пропорционально величине сигнала в децибелах, однако это сопряжено с трудностями чисто технического характера.

Временная дискретизация и амплитудное квантование сигнала неизбежно вносят в сигнал шумовые искажения, уровень которых принято оценивать по формуле $6N + 10\lg(F_{\text{дискр}}/2F_{\text{макс}}) + C$ (дБ), где константа C варьируется для разных типов сигналов: для чистой синусоиды это 1.7 дБ, для звуковых сигналов - от -15 до 2 дБ. Отсюда видно, что к снижению шумов в рабочей полосе частот $0..F_{\text{макс}}$ приводит не только увеличение разрядности отсчета, но и повышение частоты дискретизации относительно $2F_{\text{макс}}$, поскольку шумы квантования "размазываются" по всей полосе вплоть до частоты дискретизации, а звуковая информация занимает только нижнюю часть этой полосы.

В большинстве современных цифровых звуковых систем используются стандартные частоты дискретизации 44.1 и 48 кГц, однако частотный диапазон сигнала обычно ограничивается возле 20 кГц для оставления запаса по отношению к теоретическому пределу. Также наиболее распространено 16-разрядное квантование по уровню, что дает предельное соотношение сигнал/шум около 98 дБ. В студийной аппаратуре используются более высокие разрешения - 18-, 20- и 24-разрядное квантование при частотах дискретизации 56, 96 и 192 кГц. Это делается для того, чтобы сохранить высшие гармоники звукового сигнала, которые непосредственно не воспринимаются слухом, но влияют на формирование общей звуковой картины.

Для оцифровки более узкополосных и менее качественных сигналов частота и разрядность дискретизации могут снижаться; например, в телефонных линиях применяется 7- или 8-разрядная оцифровка с частотами 8..12 кГц.

Представление аналогового сигнала в цифровом виде называется также импульсно-кодовой модуляцией (ИКМ, РСМ - Pulse Code Modulation), так как сигнал представляется в виде серии импульсов постоянной частоты (временная дискретизация), амплитуда которых передается цифровым кодом (амплитудная дискретизация). РСМ-поток может быть как параллельным, когда все биты каждого отсчета передаются одновременно по нескольким линиям с частотой дискретизации, так и последовательным, когда биты передаются друг за другом с более высокой частотой по одной линии.

Сам цифровой звук и относящиеся к нему вещи принято обозначать общим термином Digital Audio; аналоговая и цифровая части звуковой системы обозначаются терминами Analog Domain и Digital Domain.

Что такое АЦП и ЦАП?

Аналогово-цифровой и цифро-аналоговый преобразователи. Первый преобразует аналоговый сигнал в цифровое значение амплитуды, второй выполняет обратное преобразование. В англоязычной литературе применяются термины ADC и DAC, а совмещенный преобразователь называют codec (coder-decoder).

Принцип работы АЦП состоит в измерении уровня входного сигнала и выдаче результата в цифровой форме. В результате работы АЦП непрерывный аналоговый сигнал превращается в импульсный, с одновременным измерением амплитуды каждого импульса. ЦАП получает на входе цифровое значение амплитуды и выдает на выходе импульсы напряжения или тока нужной величины, которые расположенный за ним интегратор (аналоговый фильтр) превращает в непрерывный аналоговый сигнал.

Для правильной работы АЦП входной сигнал не должен изменяться в течение времени преобразования, для чего на его входе обычно помещается схема выборки-хранения, фиксирующая мгновенный уровень сигнала и сохраняющая его в течение всего времени преобразования. На выходе ЦАП также может устанавливаться подобная схема, подавляющая влияние переходных процессов внутри ЦАП на параметры выходного сигнала.

При временной дискретизации спектр полученного импульсного сигнала в своей нижней части $0..F_a$ повторяет спектр исходного сигнала, а выше содержит ряд отражений (aliases, зеркальных спектров), которые расположены вокруг частоты дискретизации F_d и ее гармоник (боковые полосы). При этом первое отражение спектра от частоты F_d в случае $F_d = 2F_a$ располагается непосредственно за полосой исходного сигнала, и требует для его подавления аналогового фильтра (anti-alias filter) с высокой крутизной среза. В АЦП этот фильтр устанавливается на входе, чтобы исключить перекрытие спектров и их интерференцию, а в ЦАП - на выходе, чтобы подавить в выходном сигнале надтональные помехи, внесенные временной дискретизацией.

Что такое передискретизация (oversampling)?

Это дискретизация сигнала с частотой, превышающей основную частоту дискретизации. Передискретизация может быть аналоговой, когда с повышенной частотой делаются выборки исходного сигнала, или цифровой, когда между уже существующими цифровыми отсчетами вставляются дополнительные, рассчитанные путем интерполяции. Другой способ получения значений промежуточных отсчетов состоит во вставке нулей, после чего вся последовательность подвергается цифровой фильтрации. В АЦП используется аналоговая передискретизация, в ЦАП - цифровая.

Передискретизация используется для упрощения конструкций АЦП и ЦАП. По условиям задачи на входе АЦП и выходе ЦАП должен быть установлен аналоговый фильтр с АЧХ, линейной в рабочем диапазоне и круто спадающей за его пределами. Реализация такого аналогового фильтра весьма сложна; в то же время при повышении частоты дискретизации вносимые ею отражения спектра пропорционально отодвигаются от основного сигнала, и аналоговый фильтр может иметь гораздо меньшую крутизну среза.

Другое преимущество передискретизации состоит в том, что ошибки амплитудного квантования (шум дробления), распределенные по всему спектру квантуемого сигнала, при повышении частоты дискретизации распределяются по более широкой полосе частот, так что на долю основного звукового сигнала приходится меньшее количество шума. Каждое удвоение частоты снижает уровень шума квантования на 3 дБ; поскольку один двоичный разряд эквивалентен 6 дБ шума, каждое учетверение частоты позволяет уменьшить разрядность преобразователя на единицу.

Передискретизация вместе с увеличением разрядности отсчета, интерполяцией отсчетов с повышенной точностью и выводом их на ЦАП надлежащей разрядности позволяет несколько улучшить качество восстановления звукового сигнала. По этой причине даже в 16-разрядных системах нередко применяются 18- и 20-разрядные ЦАП с передискретизацией.

АЦП и ЦАП с передискретизацией за счет значительного уменьшения времени преобразования могут обходиться без схемы выборки-хранения.

Как устроены и работают АЦП и ЦАП?

В основном применяется три конструкции АЦП:

- параллельные - входной сигнал одновременно сравнивается с эталонными уровнями набором схем сравнения (компараторов), которые формируют на выходе двоичное значение. В таком АЦП количество компараторов равно $(2^N - 1)$, где N - разрядность цифрового кода (для восьмиразрядного - 255), что не позволяет наращивать разрядность свыше 10-12.
- последовательного приближения - преобразователь при помощи вспомогательного ЦАП генерирует эталонный сигнал, сравниваемый со входным. Эталонный сигнал последовательно изменяется по принципу половинного деления (дихотомии), который используется во многих методах сходящегося поиска прикладной математики. Это позволяет завершить преобразование за количество тактов, равное разрядности слова, независимо от величины входного сигнала.
- с измерением временных интервалов - широкая группа АЦП, использующая для измерения входного сигнала различные принципы преобразования уровней в пропорциональные временные интервалы, длительность которых измеряется при помощи тактового генератора высокой частоты. Иногда называются также считающими АЦП.

Среди АЦП с измерением временных интервалов преобладают следующие три типа:

- последовательного счета, или однократного интегрирования (single-slope) - в каждом такте преобразования запускается генератор линейно возрастающего напряжения, которое сравнивается со входным. Обычно такое напряжение получают на вспомогательном ЦАП, подобно АЦП последовательного приближения.
- двойного интегрирования (dual-slope) - в каждом такте преобразования входной сигнал заряжает конденсатор, который затем разряжается на источник опорного напряжения с измерением длительности разряда.
- следящие - вариант АЦП последовательного счета, при котором генератор эталонного напряжения не перезапускается в каждом такте, а изменяет его от предыдущего значения до текущего.

Наиболее популярным вариантом следящего АЦП является sigma-delta, работающий на частоте F_s , значительно (в 64 и более раз) превышающей частоту дискретизации F_d выходного цифрового сигнала. Компаратор такого АЦП выдает значения пониженной разрядности (обычно однобитовые - 0/1), сумма которых на интервале дискретизации F_d пропорциональна величине отсчета. Последовательность малоразрядных значений подвергается цифровой фильтрации и понижению частоты следования (decimation), в результате чего получается серия отсчетов с заданной разрядностью и частотой дискретизации F_d .

Для улучшения соотношения сигнал/шум и снижения влияния ошибок квантования, которое в случае однобитового преобразователя получается довольно высоким, применяется метод формовки шума (noise shaping) через схемы обратной связи по ошибке и цифрового фильтрования. В результате применения этого метода форма спектра шума меняется так, что основная шумовая энергия вытесняется в область выше половины частоты F_s , незначительная часть остается в нижней половине, и практически весь шум удаляется из полосы исходного аналогового сигнала.

ЦАП в основном строятся по трем принципам:

- взвешивающие - с суммированием взвешенных токов или напряжений, когда каждый разряд входного слова вносит соответствующий своему двоичному весу вклад в общую величину получаемого аналогового сигнала; такие ЦАП называют также параллельными или многоразрядными (multibit).
- sigma-delta, с предварительной цифровой передискретизацией и выдачей малоразрядных (обычно однобитовых) значений на схему формирования эталонного заряда, которые со столь же высокой частотой добавляются к выходному сигналу. Такие ЦАП носят также название bitstream.
- с широтно-импульсной модуляцией (ШИМ, Pulse Width Modulation, PWM), когда на схему выборки-хранения аналогового сигнала выдаются импульсы постоянной амплитуды и переменной длительности, управляя дозированием выдаваемого на выход заряда. На этом принципе работают преобразователи MASH (Multi-stage Noise Shaping - многостадийная формовка шума) фирмы Matsushita. Свое название эти ЦАП получили по причине применения в них нескольких последовательных формирователей шума.

При использовании передискретизации в десятки раз (обычно - 64х..512х) становится возможным уменьшить разрядность ЦАП без ощутимой потери качества сигнала; ЦАП с меньшим числом разрядов обладают также лучшей линейностью. В пределе количество разрядов может сокращаться до одного. Форма выходного сигнала таких ЦАП представляет собой полезный сигнал, обрамленный значительным количеством высокочастотного шума, который, тем не менее, эффективно подавляется аналоговым фильтром даже среднего качества.

ЦАП являются "прямыми" устройствами, в которых преобразование выполняется проще и быстрее, чем в АЦП, которые в большинстве своем - последовательные и более медленные устройства.

Каковы достоинства и недостатки цифрового звука?

Цифровое представление звука ценно прежде всего возможностью бесконечного хранения и тиражирования без потери качества, однако преобразование из аналоговой формы в цифровую и обратно все же неизбежно приводит к частичной его потере. Наиболее неприятные на слух искажения, вносимые на этапе оцифровки - гранулярный шум, возникающий при квантовании сигнала по уровню из-за округления амплитуды до ближайшего дискретного значения. В отличие от простого широкополосного шума, вносимого ошибками квантования, гранулярный шум представляет собой гармонические искажения сигнала, наиболее заметные в верхней части спектра.

Мощность гранулярного шума обратно пропорциональна количеству ступеней квантования, однако из-за логарифмической характеристики слуха при линейном квантовании (постоянная величина ступени) на тихие звуки приходится меньше ступеней квантования, чем на громкие, и в результате основная плотность нелинейных искажений приходится на область тихих звуков. Это приводит к ограничению динамического диапазона, который в идеале (без учета гармонических искажений) был бы равен соотношению сигнал/шум, однако необходимость ограничения этих искажений снижает динамический диапазон для 16-разрядного кодирования до 50-60 дБ. Положение могло бы спасти логарифмическое квантование, однако его реализация в реальном времени весьма сложна и дорога.

Искажения, вносимые гранулярным шумом, можно уменьшить путем добавления к сигналу обычного белого шума (случайного или псевдослучайного сигнала), амплитудой

в половину младшего значащего разряда; такая операция называется сглаживанием (dithering). Это приводит к незначительному увеличению уровня шума, зато ослабляет корреляцию ошибок квантования с высокочастотными компонентами сигнала и улучшает субъективное восприятие. Сглаживание применяется также перед округлением отсчетов при уменьшении их разрядности. По существу, dithering и noise shaping являются частными случаями одной технологии - с той разницей, что в первом случае используется белый шум с равномерным спектром, а во втором - шум со специально "формованным" спектром.

При восстановлении звука из цифровой формы в аналоговую возникает проблема сглаживания ступенчатой формы сигнала и подавления гармоник, вносимых частотой дискретизации. Из-за неидеальности АЧХ фильтров может происходить либо недостаточное подавление этих помех, либо избыточное ослабление полезных высокочастотных составляющих. Плохо подавленные гармоники частоты дискретизации искажают форму аналогового сигнала (особенно в области высоких частот), что создает впечатление "шероховатого", "грязного" звука.

Какие интерфейсы используются для передачи цифрового звука?

S/PDIF (Sony/Philips Digital Interface Format - формат цифрового интерфейса фирм Sony и Philips) - цифровой интерфейс для бытовой радиоаппаратуры.

AES/EBU (Audio Engineers Society / European Broadcast Union - общество звукоинженеров / европейское вещательное объединение) - цифровой интерфейс для студийной радиоаппаратуры.

Оба интерфейса являются последовательными и используют одинаковый формат сигнала и систему кодирования - самосинхронизирующийся код ВМС (Biphase-Mark Code - код с представлением единицы двойным изменением фазы), и могут передавать сигналы в формате РСМ разрядностью до 24 бит на частотах дискретизации до 48 кГц.

Каждый отсчет сигнала передается 32-разрядным словом, в котором 20 разрядов используются для передачи отсчета, а 12 - для формирования синхронизирующей преамбулы, передачи дополнительной информации и бита четности. 4 разряда из служебной группы могут использоваться для расширения формата отсчетов до 24 разрядов.

Помимо бита четности, служебная часть слова содержит бит достоверности (Validity), который должен быть нулевым для каждого достоверного отсчета. В случае приема слова с единичным битом Validity либо с нарушением четности в слове приемник трактует весь отсчет как ошибочный и может на выбор либо заменить его предыдущим значением, либо интерполировать на основе нескольких соседних достоверных отсчетов. Отсчеты, помеченные как недостоверные, могут передавать CD-проигрыватели, DAT-магнитофоны и другие устройства, если при считывании информации с носителя не удалось скорректировать возникшие в процессе чтения ошибки.

Стандартно формат кодирования предназначен для передачи одно- и двух-канального сигнала, однако при использовании служебных разрядов для кодирования номера канала возможна передача многоканального сигнала.

С электрической стороны S/PDIF предусматривает соединение коаксиальным кабелем с волновым сопротивлением 75 Ом и разъемами типа RCA ("тюльпан"), амплитуда сигнала

- 0.5 В. AES/EBU предусматривает соединение симметричным экранированным двухпроводным кабелем с трансформаторной развязкой по интерфейсу RS-422 с амплитудой сигнала 3-10 В, разъемы - трехконтактные типа Cannon XLR. Существуют также оптические варианты приемопередатчиков - TosLink (пластмассовое оптоволокно) и AT&T Link (стеклянное оптоволокно).

Какие методы используются для эффективного сжатия цифрового звука?

В настоящее время наиболее известны Audio MPEG, PASC и ATRAC. Все они используют так называемое "кодирование для восприятия" (perceptual coding) при котором из звукового сигнала удаляется информация, малозаметная для слуха. В результате, несмотря на изменение формы и спектра сигнала, его слуховое восприятие практически не меняется, а степень сжатия оправдывает незначительное уменьшение качества. Такое кодирование относится к методам сжатия с потерями (lossy compression), когда из сжатого сигнала уже невозможно точно восстановить исходную волновую форму.

Приемы удаления части информации базируются на особенности человеческого слуха, называемой маскированием: при наличии в спектре звука выраженных пиков (преобладающих гармоник) более слабые частотные составляющие в непосредственной близости от них слухом практически не воспринимаются (маскируются). При кодировании весь звуковой поток разбивается на мелкие кадры, каждый из которых преобразуется в спектральное представление и делится на ряд частотных полос. Внутри полос происходит определение и удаление маскируемых звуков, после чего каждый кадр подвергается адаптивному кодированию прямо в спектральной форме. Все эти операции позволяют значительно (в несколько раз) уменьшить объем данных при сохранении качества, приемлемого для большинства слушателей.

Каждый из описанных методов кодирования характеризуется скоростью битового потока (bitrate), с которой сжатая информация должна поступать в декодер при восстановлении звукового сигнала. Декодер преобразует серию сжатых мгновенных спектров сигнала в обычную цифровую волновую форму.

Audio MPEG - группа методов сжатия звука, стандартизованная MPEG (Moving Pictures Experts Group - экспертной группой по обработке движущихся изображений). Методы Audio MPEG существуют в виде нескольких типов - MPEG-1, MPEG-2 и т.д.; в настоящее время наиболее распространен тип MPEG-1.

Существует три уровня (layers) Audio MPEG-1 для сжатия стереофонических сигналов:

- 1 - коэффициент сжатия 1:4 при потоке данных 384 кбит/с;
- 2 - 1:6..1:8 при 256..192 кбит/с;
- 3 - 1:10..1:12 при 128..112 кбит/с.

Минимальная скорость потока данных в каждом уровне определяется в 32 кбит/с; указанные скорости потока позволяют сохранить качество сигнала примерно на уровне компакт-диска.

Все три уровня используют входное спектральное преобразование с разбиением кадра на 32 частотные полосы. Наиболее оптимальным в отношении объема данных и качества звука признан уровень 3 со скоростью потока 128 кбит/с и плотностью данных около 1 Мб/мин. При сжатии с более низкими скоростями начинается принудительное

ограничение полосы частот до 15-16 кГц, а также возникают фазовые искажения каналов (эффект типа фэйзера или фленжера).

Audio MPEG используется в компьютерных звуковых системах, CD-i/DVD, "звуковых" дисках CD-ROM, цифровом радио/телевидении и других системах массовой передачи звука.

PASC (Precision Adaptive Sub-band Coding - точное адаптивное внутриполосное кодирование) - частный случай Audio MPEG-1 Layer 1 со скоростью потока 384 кбит/с (сжатие 1:4). Применяется в системе DCC.

ATRAC (Adaptive TRansform Acooustic Coding - акустическое кодирование адаптивным преобразованием) базируется на стереофоническом звуковом формате с 16-разрядным квантованием и частотой дискретизации 44.1 кГц. При сжатии каждый кадр делится на 52 частотные полосы, результирующая скорость потока - 292 кбит/с (сжатие 1:5). Применяется в системе MiniDisk.

Какие форматы используются для представления цифрового звука?

Понятие формата используется в двух различных смыслах. При использовании специализированного носителя или способа записи и специальных устройств чтения/записи в понятие формата входят как физические характеристики носителя звука - размеры кассеты с магнитной лентой или диском, самой ленты или диска, способ записи, параметры сигнала, принципы кодирования и защиты от ошибок и т.п. При использовании универсального информационного носителя широкого применения - например, компьютерного гибкого или жесткого диска - под форматом понимают только способ кодирования цифрового сигнала, особенности расположения битов и слов и структуру служебной информации; вся "низкоуровневая" часть, относящаяся непосредственно к работе с носителем, в этом случае остается в ведении компьютера и его операционной системы.

Из специализированных форматов и носителей цифрового звука в настоящее время наиболее известны следующие:

- CD (Compact Disk - компакт-диск) - односторонний пластмассовый диск с оптической лазерной записью и считыванием, диаметром 120 или 90 мм, вмещающий максимум 74 минуты стереозвучания с частотой дискретизации 44.1 кГц и 16-разрядным линейным квантованием. Система предложена фирмами Sony и Philips и носит название CD-DA (Compact Disk - Digital Audio). Для защиты от ошибок используется двойной код Рида-Соломона с перекрестным перемежением (Cross Interleaved Reed-Solomon Code, CIRC) и модуляция кодом Хэмминга 8-14 (Eight-to-Fourteen Modulation, EFM). Различаются штампованные (CD) однократно записываемые (CD-R) и многократно перезаписываемые (CD-RW) компакт-диски.
- ИКМ-приставка (PCM deck) - система для преобразования цифрового звукового сигнала в псевдовидеосигнал, совместимый с популярными видеоформатами (NTSC, PAL/SECAM), и обратно. ИКМ-приставки применяются в сочетании с бытовыми (VHS) или студийными (S-VHS, Beta, U-Matic) видеомагнитофонами, используя их в качестве устройств чтения/записи. Устройства работают с 16-разрядным линейным квантованием на частотах дискретизации 44.056 кГц (NTSC) и 44.1 кГц (PAL/SECAM), и позволяют записывать двух- или четырехканальную цифровую сигналограмму. По сути, такая приставка представляет собой модем (модулятор-демодулятор) для видеосигнала.

- S-DAT (Stationary head Digital Audio Tape - цифровая звуковая лента с неподвижной головкой) - система наподобие обычного кассетного магнитофона, запись и чтение в которой ведутся блоком неподвижных тонкопленочных головок на ленте шириной 3.81 мм в двухсторонней кассете размером 86 x 55.5 x 9.5 мм. Реализует 16-разрядную запись двух или четырех каналов на частотах 32, 44.1 и 48 кГц.
- R-DAT (Rotary head Digital Audio Tape - цифровая звуковая лента с вращающейся головкой) - система наподобие видеоманитофона с поперечно-наклонной записью вращающимися головками. Наиболее популярный формат ленточной цифровой записи, системы R-DAT часто обозначаются просто DAT. В R-DAT используется кассета размером 73 x 54 x 10.5 мм, с лентой шириной 3.81 мм, а сама система кассеты и магнитофона очень похожа на типовой видеоманитон. Базовая скорость движения ленты - 8.15 мм/с, скорость вращения блока головок - 2000 об/мин. R-DAT работает с двухканальным (в ряде моделей - четырехканальным) сигналом на частотах дискретизации 44.1 и 48 кГц с 16-разрядным линейным квантованием, и 32 кГц - с 12-разрядным нелинейным. Для защиты от ошибок используется двойной код Рида-Соломона и модуляция кодом 8-10. Емкость кассеты - 80..240 минут в зависимости от скорости и длины ленты. Бытовые DAT-магнитофоны обычно оснащены системой защиты от незаконного копирования фонограмм, не допускающей записи с аналогового входа на частоте 44.1 кГц, а также прямого цифрового копирования при наличии запрещающих кодов SCMS (Serial Code Management System). Студийные магнитофоны таких ограничений не имеют.
- DASH (Digital Audio Stationary Head) - система с записью на магнитную ленту шириной 6.3 и 12.7 мм в продольном направлении неподвижными головками. Скорость движения ленты - 19.05, 38.1, 76.2 см/с. Реализует 16-разрядную запись с частотами дискретизации 44.056, 44.1 и 48 кГц от 2 до 48 каналов.
- ADAT (Alesis DAT) - собственная (proprietary) система восьмиканальной записи звука на видеокассету типа S-VHS, разработанная фирмой Alesis. Использует 16-разрядное линейное квантование на частоте 48 кГц, емкость кассеты составляет до 60 минут на каждый канал. Магнитофоны ADAT допускают каскадное соединение, в результате чего может быть собрана система 128-канальной синхронной записи. Для ADAT выпускается множество различных интерфейсных блоков для сопряжения с DAT, CD, MIDI и т.п. Модель Meridian (ADAT Type II) использует 20-разрядное квантование на частотах 44.1 и 48 кГц.
- DCC (Digital Compact Cassette - цифровая компакт-кассета) - бытовая система записи в продольном направлении на стандартную компакт-кассету, разработанная Philips. Скорость движения ленты - 4.76 см/с, максимальное время звучания такое же, как при аналоговой записи. Частоты дискретизации - 32, 44.1, 48 кГц, разрешение - 16/18 разрядов (метод сжатия PASC). На DCC-магнитофонах могут воспроизводиться (но не записываться) обычные аналоговые компакт-касеты. В настоящее время система DCC признана неперспективной.
- MD (MiniDisk) - бытовая и концертная система записи на магнитооптический диск, разработанная Sony. Диск диаметром 64 мм, помещенный в пластмассовый футляр размером 70 x 67.5 x 5 мм, вмещает 74 минуты (60 в ранних версиях) стереофонического звучания. При обмене со внешними устройствами используется формат 16-разрядных отсчетов на частоте 44.1 кГц, однако на сам диск сигнал записывается после сжатия методом ATRAC.

Из универсальных компьютерных форматов наиболее популярны следующие:

- Microsoft RIFF/WAVE (Resource Interchange File Format/Wave - формат файлов передачи ресурсов/волновая форма) - стандартный формат звуковых файлов в компьютерах IBM PC. Файл этого формата содержит заголовок, описывающий общие параметры файла, и один или более фрагментов (chunks), каждый из которых представляет собой волновую форму или вспомогательную информацию - режимы и порядок воспроизведения, пометки, названия и координаты участков волны и т.п. Файлы этого формата имеют расширение .WAV.
- Apple AIFF (Audio Interchange File Format - формат файла обмена звуком) - стандартный тип звукового файла в системах Apple Macintosh. Похож на RIFF и также позволяет размещать вместе со звуковой волной дополнительную информацию, в частности - сэмплы WaveTable-инструментов вместе с параметрами синтезатора.
- Формат "чистой оцифровки" RAW, не содержащий заголовка и представляющий собой только последовательность отсчетов звуковой волны. Обычно оцифровка хранится в 16-разрядном знаковом (signed) формате, когда первыми в каждой паре идут отсчеты левого канала, хотя могут быть и исключения.

Как выполняется обработка цифрового звука?

Цифровой звук обрабатывается посредством математических операций, применяемых к отдельным отсчетам сигнала, либо к группам отсчетов различной длины. Выполняемые математические операции могут либо имитировать работу традиционных аналоговых средств обработки (микширование двух сигналов - сложение, усиление/ослабление сигнала - умножение на константу, модуляция - умножение на функцию и т.п.), либо использовать альтернативные методы - например, разложение сигнала в спектр (ряд Фурье), коррекция отдельных частотных составляющих, затем обратная "сборка" сигнала из спектра.

Обработка цифровых сигналов подразделяется на линейную (в реальном времени, над "живым" сигналом) и нелинейную - над предварительно записанным сигналом. Линейная обработка требует достаточного быстродействия вычислительной системы (процессора); в ряде случаев невозможно совмещение требуемого быстродействия и качества, и тогда используется упрощенная обработка с пониженным качеством. Нелинейная обработка никак не ограничена во времени, поэтому для нее могут быть использованы вычислительные средства любой мощности, а время обработки, особенно с высоким качеством, может достигать нескольких минут и даже часов.

Для обработки применяются как универсальные процессоры общего назначения - Intel 8035, 8051, 80x86, Motorola 68xxx, SPARC - так и специализированные цифровые сигнальные процессоры (Digital Signal Processor, DSP) Texas Instruments TMS xxx, Motorola 56xxx, Analog Devices ADSP-xxxx и др.

Разница между универсальным процессором и DSP состоит в том, что первый ориентирован на широкий класс задач - научных, экономических, логических, игровых и т.п., и содержит большой набор команд общего назначения, в котором преобладают обычные математические и логические операции. DSP специально ориентированы на обработку сигналов и содержат наборы специфических операций - сложение с ограничением, перемножение векторов, вычисление математического ряда и т.п. Реализация даже несложной обработки звука на универсальном процессоре требует значительного быстродействия и далеко не всегда возможна в реальном времени, в то время как даже простые DSP нередко справляются в реальном времени с относительно

сложной обработкой, а мощные DSP способны выполнять качественную спектральную обработку сразу нескольких сигналов.

В силу своей специализации DSP редко применяются самостоятельно - чаще всего устройство обработки имеет универсальный процессор средней мощности для управления всем устройством, приема/передачи информации, взаимодействия с пользователем, и один или несколько DSP - собственно для обработки звукового сигнала. Например, для реализации надежной и быстрой обработки сигналов в компьютерных системах применяют специализированные платы с DSP, через которые пропускается обрабатываемый сигнал, в то время как центральному процессору компьютера остаются лишь функции управления и передачи.

Что представляет собой цифровая звуковая рабочая станция?

Digital Audio Workstation (DAW) представляет собой специализированную или универсальную компьютерную систему, способную выполнять запись, хранение, воспроизведение и обработку цифрового звука. Специализированные системы ориентированы исключительно на работу с цифровым звуком и выпускаются в законченном исполнении, допускающем лишь ограниченное расширение, либо нерасширяемые вообще. Универсальные системы представляют собой обычный персональный компьютер, снабженный средствами для ввода/вывода звука (ЦАП/АЦП и/или цифровые интерфейсы) и набором программ для его записи, воспроизведения и обработки. Кроме этого, станция может содержать и другие компоненты - например, аппаратные модули цифровой обработки, музыкальные синтезаторы, записывающие CD-приводы и т.п.

Поскольку любая компьютерная система является сильным источником высокочастотных помех, возникают определенные проблемы в достижении профессионального качества звука при использовании встроенных АЦП/ЦАП. В таких случаях предпочтительно использование внешних модулей АЦП/ЦАП, выдающих и получающих цифровую информацию в реальном времени через универсальные или собственные цифровые интерфейсы.

Какие жесткие диски используются в рабочих станциях?

Большинство специализированных рабочих станций используют для хранения звука жесткие диски с интерфейсом SCSI (Small Computer System Interface - интерфейс малых компьютерных систем), ставшие универсальным стандартом - любая популярная компьютерная система имеет возможность подключения этих дисков. Достоинствами SCSI является универсальность среди всех компьютерных систем, возможность подключения до семи устройств (любых, не только дисковых) к одному контроллеру, хороший арбитраж при конкуренции устройств, интеллектуальность каждого устройства, более высокое общее качество исполнения, возможность использования интерфейса для прямой связи между двумя станциями. К недостаткам SCSI следует отнести высокую стоимость интерфейсов и дисков и ограниченный спектр выпускаемых моделей.

В компьютерах типа IBM PC более популярны жесткие диски с интерфейсом IDE (Integrated Drive Electronics - электроника, встроенная в накопитель), не получившие распространения в других системах. Достоинства IDE-дисков - простота, хорошая производительность, не уступающая большинству SCSI-дисков, а в ряде случаев - превосходящая их, низкая стоимость, массовый выпуск, широкий спектр моделей. Недостатки - низкая производительность и надежность моделей низших классов,

возможность подключения только двух накопителей к одному контроллеру, невозможность прямого соединения двух станций, часто худшая поддержка драйверами операционных систем.

Есть ли преимущества у SCSI перед IDE при цифровой записи?

Среди пользователей звуковых рабочих станций - как домашних, так и студийных - бытует мнение, что только диски SCSI способны обеспечить нужное быстродействие. Однако, несмотря на ряд очевидных преимуществ SCSI, большинство даже профессиональных рабочих станций на IBM PC вполне может обходиться дисками IDE. Скорость чтения/записи типовых моделей IDE-дисков сегодня (конец 1998 г.) находится на уровне 6-10 Мб/с при времени поиска около 8-10 мс, что равнозначно таким же типовым (не High End) моделям SCSI.

Такой жесткий диск свободно справляется с одновременным чтением 16-разрядных звуковых данных по 20-30 звуковым каналам на частоте дискретизации 48 кГц, и несколько меньшим объемом данных в случае записи. Другое дело, что в случае SCSI его внутренняя оптимизация (сортировка запросов для минимизации перемещения головок в SCSI-2) часто маскирует неоптимальную работу ОС и звуковой программы, а для достижения такого уровня на IDE может потребоваться хороший драйвер ОС и аккуратно сделанная программа (например, DDClip).

Причины нелюбви многих пользователей к IDE-дискам происходят оттого, что с этими дисками они обычно сталкиваются в дешевых, некачественно собранных и протестированных компьютерах средней мощности, состоящих из разномастных компонент, нередко плохо совместимых друг с другом. И напротив - SCSI-диски чаще всего ставятся в более мощные и дорогие модели, содержащие компоненты "уважаемых" производителей, более тщательно собранные и проверенные. Замена во втором варианте диска SCSI на IDE примерно равной производительности и сборка/настройка системы с учетом особенностей IDE во многих случаях не окажет заметного влияния на ее производительность.

Зачем в цифровой записи применяются жесткие диски класса AV?

Класс AV (Audio/Video) у жестких дисков означает их способность предельно равномерно, без пауз, записывать и считывать потоки данных. Такие диски снабжаются внутренним буфером большего размера и не прерывают процесса чтения/записи термокалибровкой системы позиционирования. Для систем цифровой записи, имеющих недостаточное быстродействие и объемы ОЗУ, чтобы сгладить возможные неравномерности в работе обычных дисков, диски класса AV являются единственным возможным выходом.

Следует иметь в виду, что наличие аббревиатуры AV в обозначении диска еще не означает его принадлежности к классу Audio/Video - об этом должно быть явно упомянуто в паспорте диска.

Однако указанная особенность в общем случае необходима только при работе с качественной видеoinформацией, скорость поступления которой составляет порядка 10 мегабайт в секунду на канал. В случае же звуковых систем скорость одноканального 16-разрядного потока с частотой дискретизации 48 кГц на два порядка меньше и составляет всего 94 килобайта в секунду. В то же время почти никакая рабочая станция не в состоянии обеспечить одновременную работу с сотней каналов, как и жесткий диск не в

состоянии параллельно обрабатывать такое количество данных, расположенных в разных его участках. В реальных применениях многоканальной записи на одном диске основная часть накладных расходов дисковой подсистемы ложится на перемещение головок между участками записи, а отнюдь не на саму передачу данных. Низкая же скорость звуковых потоков делает более удобной и надежной их буферизацию в ОЗУ компьютера, компенсирующую термокалибровку диска в течение 0.5 - 1 с, нежели использование дорогих и редких дисков AV-класса. К тому же далеко не на всех обычных дисках термокалибровка оказывает заметное влияние на равномерность потока данных.

"Рваная" передача данных может также возникать при использовании "неправильной" операционной системы (DOS, Windows без 32-разрядного драйвера диска и т.п.), недостаточном количестве и размере файловых буферов ОС и записывающей программы, применении дисков низкого класса со скоростью передачи порядка 1-2 мегабайт в секунду и ниже, неправильном подключении диска и т.п. В любом случае, такие ситуации чаще всего говорят о неправильной конфигурации и настройке аппаратной и программной части системы.

Искажается ли цифровой сигнал при передаче и хранении?

Поскольку любой цифровой сигнал представляется реальной электрической кривой напряжения или тока - его форма так или иначе искажается при любой передаче, а "замороженный" для хранения сигнал (сигналограмма) подвержен деградации в силу обычных физических причин. Все эти воздействия на форму несущего сигнала являются помехами, которые до определенной величины не изменяют информационного содержания сигнала, как отдельные искажения и выпадения букв в словах обычно не мешают правильному пониманию этих слов, причем избыточность информации, как и увеличение длины слов, повышает вероятность успешного распознавания. Другими словами, сам несущий сигнал может искажаться, однако переносимая им информация - закодированный звуковой сигнал - в абсолютном большинстве случаев остается неизменной.

Для того, чтобы качество несущего сигнала не ухудшалось, любая передача полезной звуковой информации - копирование, запись на носитель и считывание с него - обязательно должна включать операцию восстановления формы несущего сигнала, а в идеале - и первичного цифрового вида сигнала информационного, и лишь после этого заново сформированный несущий сигнал может быть передан следующему потребителю. В случае прямого копирования без восстановления (например, обычным переписыванием видеокассеты с цифровым сигналом, полученным при помощи ИКМ-приставки, на обычных видеомагнитофонах) качество цифрового сигнала ухудшается, хотя он по-прежнему полностью содержит всю переносимую им информацию. Однако после многократного последовательного копирования или длительного хранения качество ухудшается настолько, что начинают возникать неисправимые ошибки, необратимо искажающие переносимую сигналом информацию. Поэтому копирование и передачу цифровых сигналов необходимо вести только в цифровых устройствах, а при хранении на носителях - своевременно "освежать" не дожидаясь необратимой деградации (для магнитных носителей этот срок оценивается в несколько лет). Правильно переданная или обновленная цифровая сигналограмма качества не теряет и может копироваться и существовать вечно в абсолютно неизменном виде.

Тем не менее, не следует забывать, что корректирующая способность любого кода конечна, а реальные носители далеки от идеальных, поэтому возникновение неисправимых ошибок - на такая уж редкая вещь, особенно при неаккуратном обращении

с носителем. При чтении с новых и правильно хранимых DAT-кассет или компакт-дисков в качественных и надежных аппаратах таких ошибок практически не возникает, однако при старении, загрязнении и повреждении носителей и считывающих систем их становится больше. Одиночная неисправленная ошибка почти всегда незаметна на слух благодаря интерполяции, однако она приводит к искажению формы исходного звукового сигнала, а накопление таких ошибок со временем начинает ощущаться и на слух.

Отдельную проблему составляет сложность регистрации неисправленных ошибок, а также проверки идентичности оригинала и копии. Чаще всего конструкторы цифровых звуковых устройств, работающих в реальном времени, не озабочены вопросом точной проверки достоверности передачи, считая вполне достаточными меры, принятые для коррекции ошибок. Невозможность в общем случае повторной передачи ошибочного отсчета или блока приводит к тому, что интерполяция происходит скрытно и после копирования нельзя с уверенностью сказать, точно ли скопирован исходный сигнал. Индикаторы ошибки, имеющиеся в ряде устройств, обычно включаются только в момент ее возникновения, и в случае одиночных ошибок их срабатывание легко может остаться незамеченным. Даже в системах на основе персональных компьютеров чаще всего нет возможности контролировать правильность приема по цифровому интерфейсу или прямого считывания CD; выходом является только многократное повторение операции и сравнение результатов.

И наконец, в принципе возможны ситуации, когда даже незначительные ошибки способны необратимо исказить передаваемую информацию, оставшись при этом незамеченными системой передачи. Другое дело, что вероятность возникновения подобных ошибок исчезающе мала (порядка одной на несколько лет непрерывной передачи сигнала), поэтому такую возможность практически нигде не принимают в расчет.

Страдает ли качество сигнала при цифровой обработке?

Прежде всего, необходимо различать "искажающие" и "неискажающие" виды обработки. К первым относятся операции, изменяющие форму и структуру сигнала - смешивание, усиление, фильтрация, модуляция и т.п., ко вторым - операции монтажа (вырезка, вклейка, наложение) и переноса (копирования).

Качество сигнала может страдать только при "искажающей" обработке, причем любой - и аналоговой, и цифровой. В первом случае это происходит в результате внесения шумов, гармонических, интермодуляционных и других искажений в узлах аналогового тракта, во втором - благодаря конечной точности квантования сигнала и математических вычислений. Все цифровые вычисления выполняются в некоторой разрядной сетке фиксированной длины - 16, 20, 24, 32, 64, 80 и более бит; увеличение разрядности сетки повышает точность вычислений и уменьшает ошибки округления, однако в общем случае не может исключить их полностью. Конечная точность квантования первичного аналогового сигнала приводит к тому, что даже при абсолютно точной обработке полученного цифрового сигнала квантованное значение каждого отсчета все равно отличается от своего идеального значения. Для минимизации искажений при обработке в студиях предпочитают обрабатывать и хранить сигналограммы на мастер-носителях с повышенным разрешением (20, 24 или 32 разряда), даже если результат будет тиражироваться на носителе с меньшим разрешением.

Кроме собственно ошибок вычислений и округления, на точность сильно влияет выбор представления числовых отсчетов сигнала при обработке. Традиционное представление PCM с так называемой фиксированной точкой (fixed point), когда отсчеты представляются

целыми числами, наиболее удобно и влечет минимум накладных расходов, однако точность вычислений зависит от масштаба операций - например, при умножении образуются числа вдвое большей разрядности, которые потом приходится приводить обратно к разрядности исходных отсчетов, а это может привести к переполнению разрядной сетки. Компромиссным вариантом служит промежуточное увеличение разрядности отсчетов (например, 16->32), что снижает вероятность переполнения, однако требует большей вычислительной мощности, объема памяти и вносит дополнительные искажения при обратном понижении разрядности. Кроме того, снижению погрешности способствует правильный выбор последовательности коммутативных (допускающих перестановку) операций, группировка дистрибутивных операций, учет особенностей работы конкретного процессора и т.п.

Другим способом увеличения точности является преобразование отсчетов в форму с плавающей точкой (floating point) с разделением на значащую часть - мантиссу и показатель величины - порядок. В этой форме все операции сохраняют разрядность значащей части, и умножение не приводит к переполнению разрядной сетки. Однако, как само преобразование между формами с фиксированной и плавающей точкой, так и вычисления в этой форме требуют на порядки большего быстродействия процессора, что сильно затрудняет их использование в реальном времени.

Несмотря на то, что качество сигнала неизбежно, хоть и незначительно, ухудшается при любой "искажающей" цифровой обработке, некоторые операции при определенных условиях являются полностью и однозначно обратимыми. Например, усиление сигнала по амплитуде в три раза заключается в умножении каждого отсчета на три; если эта операция выполнялась с фиксированной точкой и при этом не возникло переполнения, с помощью деления на три потом можно будет вернуть все отсчеты в исходное состояние, тем самым полностью восстановив первоначальное состояние сигнала. И в то же время после умножения каждый отсчет окажется увеличенным точно в три раза, поэтому ошибка относительно исходного аналогового сигнала, внесенная при квантовании, также увеличится в среднем в три раза, тем самым ухудшив общее качество сигнала.

Сказанное выше демонстрирует, что ухудшение качества при "искажающей" цифровой обработке совсем не обязательно накапливается со временем, хотя в большинстве реальных применений происходит именно так. Кроме того, это не означает, что любая операция цифрового усиления всегда будет однозначно обратимой - это зависит от многих особенностей применения операции. Тем не менее, грамотно и качественно реализованная цифровая обработка может давать существенно меньший уровень искажений, чем такая же аналоговая, разве что это будут искажения разных видов.

Страдает ли качество сигнала при преобразованиях форматов?

Только в том случае, когда в процессе преобразования применяются "искажающие" операции - изменение разрядности отсчета, частоты дискретизации, фильтрация, сжатие с потерями и т.п. Простое увеличение разрядности отсчета с сохранением частоты дискретизации будет неискажающим, однако такое же увеличение, сопряженное с применением сглаживающей функции - уже нет. Уменьшение разрядности отсчета всегда является искажающей операцией, кроме случая, когда преобразуемые отсчеты были получены таким же простым увеличением разрядности - равной или меньшей.

Многие форматы отличаются друг от друга только порядком битов в слове, отсчетов левого и правого каналов в потоке и служебной информацией - заголовками, контрольными суммами, помехозащитными кодами и т.п. Точный способ проверки

неискажаемости сигнала заключается в преобразовании нескольких различных потоков (файлов) формата F1 в формат F2, а затем обратно в F1. Если информационная часть каждого потока (файла) при этом будет идентична исходной - данный вид преобразования можно считать неискажающим.

Под информационной частью потока (файла) понимается собственно набор данных, описывающих звуковой сигнал; остальная часть считается служебной и на форму сигнала в общем случае не влияет. Например, если в служебной части файла или потока предусмотрено поле для времени его создания (передачи), то даже в случае полного совпадения информационных частей двух разных файлов или потоков их служебные части окажутся различными, и это будет зафиксировано логическим анализатором в случае потока или программой побайтного сравнения - в случае файла. Кроме этого, временной сдвиг одного сигнала относительно другого, возникающий при выравнивании цифрового потока по границам слов или блоков и состоящий в добавлении нулевых отсчетов в начало и/или конец файла или потока, также приводит к их кажущемуся цифровому несовпадению. В таких ситуациях для проверки идентичности цифровых сигналов необходимо пользоваться специальной аппаратурой или программой.

Как преобразовать цифровой звук из одного формата в другой?

Для "перегонки" звука между специализированными системами, имеющими совместимые цифровые интерфейсы, достаточно соединить их цифровым кабелем и переписать звук с одной системы на другую; в ряде сочетаний устройств при этом возможно ухудшение качества сигнала из-за уменьшения разрядности отсчета, передискретизации или сжатия звука. Например, при копировании звука между одинаковыми системами MiniDisk через интерфейс S/PDIF сжатый звуковой поток на передающей стороне подвергается восстановлению, а на приемной - повторному сжатию. Вследствие несимметричности алгоритма ATRAC в звук при повторном сжатии будут внесены добавочные искажения.

Для преобразования компьютерного файла в другой формат используются программы-конверторы: WAV2AIFF/AIFF2WAV, Convert, AWave и другие - на IBM PC, SoundExtractor, SampleEditor, BST - на Apple Macintosh.

Обмен звуковой информацией между компьютерной и специализированной системой нередко возможен несколькими способами:

- Прямой перенос по цифровому интерфейсу, если у обеих систем имеются совместимые цифровые интерфейсы. При этом на компьютерной системе используется программа записи/воспроизведения, формирующая или воспроизводящая стандартный для данной системы звуковой файл.
- Чтение/запись на специализированных системах стандартных компьютерных носителей. Например, ряд музыкальных рабочих станций использует гибкие диски в форматах стандартных файловых систем IBM PC или Macintosh, либо позволяет прочитать или создать такой диск.
- Чтение и запись на компьютерной системе специализированных носителей и их специальных форматов, если это позволяет аппаратура и программное обеспечение. Таким образом читаются и пишутся дискеты от Ensoniq, AKAI, Emulator, компакт-диски ряда "чужих" систем, а также читаются и пишутся обычные звуковые компакт-диски.

Какие компьютерные программы используются для обработки звука?

На IBM PC наиболее популярны редакторы Cool Edit Pro (Syntrillium) Sound Forge (Sonic Foundry), WaveLab (Steinberg) и системы многодорожечной записи SAW Plus, Samplitude, N-Track и DDClip. На Apple Macintosh используются программ Alchemy, Deck II, DigiTracks, HyperPrism.

Какие платы применяются для работы с цифровым звуком на IBM PC?

Для любительской работы со звуком, не требующей высокого качества, в принципе достаточно любой звуковой карты, качество которой удовлетворяет ее владельца. Самые дешевые карты типа Edison, Sky Rocket, Media Vision, Sound Blaster Vibra и т.п. обладают достаточно низким качеством преобразования звука. Более хорошим звучанием из простых карт обладают некоторые модели Sound Blaster AWE32, Gravis Ultrasound, Turtle Beach Tropez и некоторые другие. Все эти карты ориентированы на шину ISA и работают с 16-разрядным звуком.

Минимальным уровнем карт, пригодных для более-менее серьезной работы со звуком, принято считать снятую с производства Turtle Beach Tahiti (16 разрядов, 18-разрядный ЦАП) и ее выпускаемый ныне улучшенный вариант Fiji (20 разрядов). Эти карты также сделаны под ISA. Для Fiji существует отдельная дочерняя плата электрического интерфейса S/PDIF.

Представляет интерес 20-разрядная карта Terratec EWS64XL, АЦП и ЦАП которой обеспечивают 16- и 18-разрядную точность, а электрический и оптический интерфейс S/PDIF - 20-разрядную.

Более высокий класс ISA-карт представлен серией 24-разрядных карт AdB Multi!Wav с цифровыми интерфейсами S/PDIF и AES/EBU и синхронизацией Word Clock: Digital Pro18 (18-разрядный мониторный ЦАП), Digital Pro24 (24-разрядный мониторный ЦАП), Analog Pro24 (24-разрядные ЦАП и АЦП, без цифровых интерфейсов). 20-разрядная карта Zefiro Acoustics ZA-2 имеет электрические и оптические интерфейсы S/PDIF и AES/EBU, 24-разрядный DSP и мониторный ЦАП.

Спектр карт для шины PCI открывает давно известная, хоть и устаревшая, AudioMedia III (24 разряда, четыре 18-разрядных АЦП/ЦАП, вход/выход S/PDIF, 24-разрядная обработка в DSP). В последнее время популярны три 24-разрядные карты Event Electronics с 20-разрядными ЦАП/АЦП и 24-разрядным DSP: Darla (2 АЦП, 8 ЦАП), Gina (2 АЦП, 8 ЦАП, S/PDIF) и Layla (8 АЦП и 10 ЦАП в выносном модуле, S/PDIF, MIDI, Word Clock).

Связана ли разрядность отсчета с разрядностью канала передачи?

Никоим образом. В сущности, почти все современные АЦП и ЦАП - и 16-, и 20-, и 24-разрядные - работают с последовательными потоками однобитовых данных, точно так же передается информация в большинстве цифровых интерфейсов и хранится на цифровых носителях. При этом частота следования отдельных битов, составляющих отсчет, повышается настолько, чтобы обеспечить передачу нужного количества битов в течение интервала дискретизации, что полностью эквивалентно параллельной передаче отсчетов непосредственно с частотой дискретизации.

Также в ряде систем используется последовательно-параллельная передача, когда отсчеты передаются группами битов меньшей длины (обычно байтами). Например, большинство 16-разрядных компьютерных звуковых карт использует для передачи отсчетов 8-разрядные каналы прямого доступа к памяти (DMA), где каждый 16-разрядный

отсчет передается последовательно из двух байтов. Пропускной способности DMA достаточно для одновременной передачи более десяти стереофонических 16-разрядных звуковых потоков с частотой дискретизации 44.1 кГц, так что такие карты ничуть не уступают моделям, использующим 16-разрядные каналы, а разница в качестве звука обусловлена сугубо параметрами АЦП и ЦАП.

Что такое джиттер?

Jitter - дрожание (быстрые колебания) фазы синхросигналов в цифровых системах, приводящее к неравномерности во времени моментов срабатывания тактируемых этими сигналами цифровых устройств. Сами по себе цифровые устройства нечувствительны к таким колебаниям, пока они не достигают значительной величины по сравнению с общей длительностью импульсов, однако в "пограничных" устройствах, находящихся на стыке цифровой и аналоговой частей схемы - АЦП и ЦАП - джиттер приводит к неравномерности моментов срабатывания компараторов АЦП или ключей ЦАП, приводящей к нарушению правильности формы аналогового сигнала. Для высокочастотных компонент сигнала дрожание фазы приводит к "размыванию" звука - нарушению субъективной пространственной локализации источников, поскольку слуховое восприятие локализации базируется в основном на фазовых, а не на амплитудных соотношениях стереоканалов.

Джиттер может возникать из-за любой нестабильности напряжений и токов в области ЦАП/АЦП. Например, колебания питающих напряжений изменяют частоту опорного генератора, наводки на провода и печатные дорожки искажают форму цифровых сигналов. Даже если эти искажения не изменяют информационного содержимого сигнала - заключенной в нем битовой последовательности, они могут нарушить равномерность опроса входного звукового сигнала в АЦП или выдачу выходного сигнала с ЦАП и привести к искажениям формы, особенно заметной в области высоких частот.

Величина джиттера обозначает максимальное абсолютное отклонение момента перехода тактового сигнала из одного состояния в другое от расчетного значения, и измеряется в секундах. Для систем среднего качества допустимая величина джиттера составляет порядка 100 пикосекунд, для систем класса Hi-Fi ее стараются предельно минимизировать.

Для борьбы с джиттером используется тактирование АЦП и ЦАП высокостабильными генераторами, а для подавления неравномерности цифрового потока, поступающего на ЦАП - промежуточными буферами типа FIFO (очередь). Для уменьшения влияния помех применяются обычные методы - экранирование, развязки, исключение "земляных петель", отдельные источники питания, питание критичных схем от аккумулятора и т.п. Хорошие результаты дают внешние модули ЦАП, в которых реализованы описанные методы - например, Audio Alchemy DAC-in-the-Box и другие.

Необходимо различать "пограничный" джиттер, действующий на границах аналоговой и цифровой части схемы - в области АЦП или ЦАП, и "внутренний", возникающий в любых других участках чисто цифровой схемы. Влияние на звуковой сигнал имеет только "пограничный" джиттер, ибо только он непосредственно связан с преобразованием аналогового звукового сигнала. Весь "внутренний" джиттер при грамотном построении схемы должен полностью подавляться в интерфейсных цепях, однако некорректная реализация может пропускать его и непосредственно на ЦАП/АЦП.

Может ли джиттер передаваться при копировании?

Возникающий в цепях формирования, обработки, передачи, записи и чтения цифровых сигналов "внутренний" джиттер вполне может распространяться по системе, выходить за ее пределы и переноситься между системами через цифровые интерфейсы передачи или цифровые же носители информации. При этом величина джиттера может как ослабляться, так и усиливаться. При использовании интерфейсов передачи со "встроенным" (embedded) синхросигналом, а также при чтении с любого носителя, приемная сторона вынуждена синхронизироваться с передатчиком путем использования систем фазовой автоподстройки частоты (ФАПЧ, Phase Locked Loop - PLL), которая вносит дополнительные дрожания, будучи не в состоянии мгновенно отслеживать изменения фазы и частоты принимаемого сигнала.

Один из возможных способов ослабления джиттера при передаче - использование синхронных интерфейсов с отдельным тактовым сигналом (Word Clock), а еще лучше - асинхронных двунаправленных с возможностью согласования темпа передачи, наподобие RS-232. В этом случае стороны могут не опасаться возможного опустения или переполнения буфера на приемном конце, передача может выполняться блоками с более высокой скоростью, чем идет вывод звука, а приемная сторона может использовать полностью независимый стабильный генератор для извлечения отсчетов из буфера. Однако все это имеет смысл только в том случае, когда приемник работает непосредственно на ЦАП - при записи на носитель неравномерности такой величины влияния на качество звука не оказывают.

Таким образом, в корректно реализованной системе все виды джиттера, возникающие в чисто цифровых блоках и между ними, являются "внутренними" и должны быть подавлены до передачи цифрового сигнала на ЦАП для окончательного преобразования. Это может быть сделано при помощи промежуточного буфера, схемы ФАПЧ с плавным изменением частоты генератора (медленное изменение в небольших пределах, в отличие от дрожания, практически не ощущается на слух), или каким-либо другим методом.

Может ли один и тот же цифровой сигнал звучать по-разному?

Для слуховой оценки звукового сигнала его необходимо воспроизвести либо одновременно на двух разных системах, либо последовательно - на одной. Даже если в обоих случаях сам цифровой сигнал будет одинаковым, набор сопутствующих условий - аппарат, носитель, его микроструктура, первичные сигналы при считывании информации, особенности работы декодеров, спектр аналоговых шумов и помех - почти всегда будет различен. Все эти побочные процессы могут создавать паразитные наводки, искажающие форму цифрового сигнала, порождающие джиттер, воздействующие на цепи питания и прочие аналоговые компоненты системы. В правильно сконструированных и тщательно выполненных аппаратах все эти влияния должны быть подавлены до уровня, **недоступного** восприятию, однако для большинства бытовых и особенно бюджетных аппаратов это не так.

Могут быть и более прозаичные причины для возникновения разницы - такие, как неустойчивое считывание цифрового носителя, при котором декодер не в состоянии однозначно восстановить закодированный звуковой сигнал и вынужден прибегать к его интерполяции, ухудшающей качество звучания. Такая же интерполяция или гашение отсчетов происходит в случае ошибочного их приема по цифровым межсистемным интерфейсам, что может быть вызвано плохим качеством или чрезмерной длиной кабеля, воздействием на него сильных помех, неисправностью приемника или передатчика, плохой их совместимостью и т.п. Поэтому вопрос о сравнении звучания должен рассматриваться только после того, как доказана идентичность цифровых потоков,

поступающих на окончательный ЦАП. Под ЦАП здесь должен пониматься именно неделимый, "самый последний" преобразователь, а не произвольное сложное устройство, получающее на входе цифровой сигнал и выдающее на выходе аналоговый.