

УДК 004.522

АНАЛИЗ МЕТОДОВ И АЛГОРИТМОВ ОЦИФРОВОЫВАНИЯ ЗВУКОВЫХ ДАННЫХ

Иванченко Юрий Сергеевич, Завадская Татьяна Владимировна
Донецкий национальный технический университет,
г. Донецк, Донецкая народная республика

Аннотация

В рассматриваемой статье проводится анализ видов и методов преобразования аналогового звука в цифровую форму. Приведены различные инструменты реализации оцифровывания на программном и аппаратном уровнях, с подробным описанием и иллюстрацией их работы. Рассмотрены требования и условия для реализации оцифровывания, а также типы алгоритмов, с определением их преимуществ и недостатков. В заключении подведены итоги проведенных исследований, а также приведена эффективная комбинация алгоритма и метода. Результаты оцифровывания звука, в данной статье, могут использоваться для решения задач в направлении распознавания речи.

Ключевые слова: оцифровывание, частотные фильтры, аппаратная реализация, микропрограммная реализация.

METHODS AND ALGORITHMS FOR DIGITALIZATION OF SOUND DATA

Ivanchenko Yuri, Tatiana Zavadskaya
Donetsk national technical university,
Donetsk, Donetsk People's Republic

Abstract

This article analyzes the types and methods of converting analog sound into digital form. Methods for the implementation of digitization at the software and hardware levels are presented, with a detailed description and illustration of their work. The requirements and conditions for the implementation of digitization are considered, as well as the type of algorithms, using their advantages and disadvantages. In the conclusion, the results of the research carried out are summarized, and an effective combination of the algorithm and method is presented. The results of sound digitization, this article, Tasks in the direction of speech recognition.

Keywords: digitalization, frequency filters, hardware implementation, firmware implementation.

Введение

Звук – это явление, которое окружает человека всю его жизнь. С точки зрения физики – это механические колебания, в виде упругих волн, в любой среде, которые характеризуется частотой и амплитудой, или говоря иначе это волнообразное движение воздуха [3]. С иной точки зрения звук является способом для общения, взаимодействия живого существа в течении существования в социуме или природе, в целом.

Сигнал можно записывать, обрабатывать или воспроизводить. Используется аналоговый и цифровой способы передачи звуковых сигналов. Преимущества аналогового способа заключаются в простоте использования и дешевизне. Недостатки состоят в сложности хранения аналогового сигнала на существующих устройствах хранения (виниловые пластинки, кассеты) [1].

Целью исследования является определение свойств и типизирование методов и алгоритмов оцифровывания звуковых данных. Так как цифровые данные, наиболее актуальны во многих аспектах современной жизни человека, то существует необходимость быстрого, качественного и эффективного преобразования аналогового сигнала, в такой вид, который

человек не способен распознать, без устройства воспроизведения сигнала и не должен заметить разницы между преобразованиями до и после. Также хранение данных в преобразованном формате подразумевает большой срок сохранности в первоначальном виде.

В ходе данного исследования необходимо применение дискретизации для преобразования исходного сигнала в цифровую форму. Поэтому предстоит рассмотреть способы оцифровывания подробнее.

Аппаратный способ может оказаться громоздким, при этом не гарантирует высокого качества обработки. Это объясняется наличием сразу нескольких устройств и зависимости качества от цены за аппаратное устройство, например:

- устройство цифровой обработки сигнала – преобразует входной сигнал в цифровой вид или наоборот. Пренебрежение его качеством повышает вероятность на выходе шумов;
- интерфейсное устройство – реализует обмен или отправку данных между интерфейсами;
- синтезатор – данное устройство используется для наложения различных звуковых эффектов на сигнал. Его составляет набор синтезов частотной модуляции, волновой таблицы или с применением обоих сразу;
- микшер – это устройство регулирует уровни общего сигнала, а также входных сигналов (линейного или цифрового) и тембра;

Так как вышеописанные блоки аппаратных устройств воздействуют на уровень изменения спектра, амплитуды и фазы, что и влияет на конечное качество звука. Это также говорит о том, что, чем больше элементов обработки будет использовано, тем выше будет стоимость.

Программный способ должен выполнять свои задачи при малых ресурсах, что требует большой вычислительной мощности устройства и может отразиться на его стоимости. Весомый недостаток способа заключается во времени срабатывания, оно больше, чем на аналоговом устройстве. Несмотря на этот недостаток, оцифровать сигнал, размером с обычный музыкальный трек, устройство способно за доли секунды, что человек фактически не заметит. Но, в остальных случаях, важно достигнуть максимально высокой скорости оцифровки за малый временной отрезок, при любой размере обрабатываемых данных. Архитектура рассматриваемого способа состоит из следующих модулей:

- плагин – расширяет функциональные возможности преобразования речи из аналогового вида в цифровой;
- секвенсор – обеспечивает редактирование и создание аудио эффектов, на midi – синтезаторе.
- редактор цифрового аудио – реализует запись и сохранение обработанных данных на накопительном, запоминающем устройстве;
- FTP-анализатор – отвечает за измерительный анализ звуковых данных. Данные анализаторы могут строить амплитудно-частотная характеристику (АЧХ), график спектрально-временного представления звука и т.п.;

Использование программного способа имеет преимущество в размере стоимости реализации задач преобразования, отсутствии громоздких составляющих при том, что незначительно уступает в скорости срабатывания устройства; что говорит о преимуществе над аппаратным способом и выгоде использования именно программного.

Методы оцифровывания сигнала

После подробного теоретического описания устройств, которые преобразуют звуковые данные, можно перейти к анализу методик их работы и инструментов, с помощью которых стало возможно осуществить качественное преобразование аналогового сигнала в цифровой.

Практическая реализация производится с использованием фильтров [1].

Частотные фильтры подразделяются на [2]:

- высокочастотные – пропускают сигналы, которые выше заданной, пороговой частоты (рис. 1);

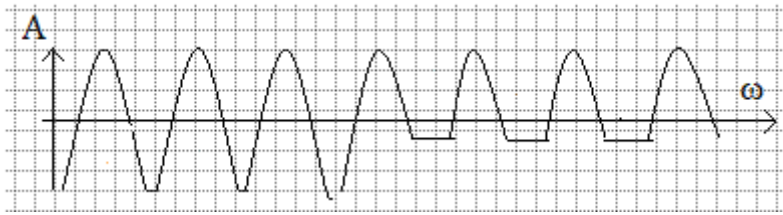


Рисунок 1 – График, иллюстрирующий работу высокочастотного фильтра

- низкочастотные – пропускают сигналы, которые ниже заданной, пороговой частоты. (ФНЧ) (рис. 2);

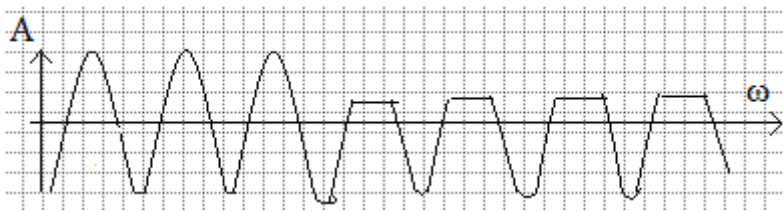


Рисунок 2 – График, иллюстрирующий работу низкочастотного фильтра

- Полосовые фильтры делятся на широкополосные, узкополосные [3] (рис. 3).

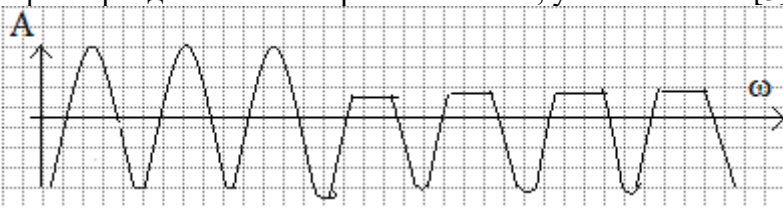


Рисунок 3 – График, иллюстрирующий работу полосовых фильтров

Шельфовые фильтры подразделяются на:

- низкочастотный шельфовые фильтры – усиливают или ослабляют сигналы, ниже определённой частоты (рис. 4).

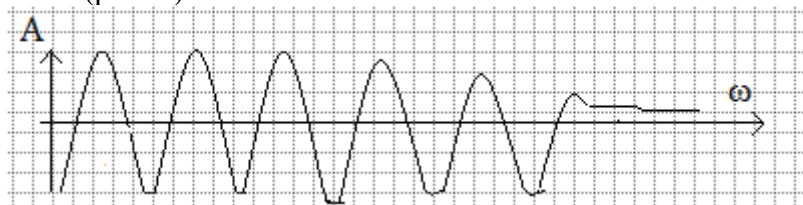


Рисунок 4 – График, иллюстрирующий работу низкочастотных шельфовых фильтров, в режиме ослабления сигнала

- высокочастотный шельфовый фильтр – усиливают или ослабляют сигналы, выше определённой частоты (рис. 5).

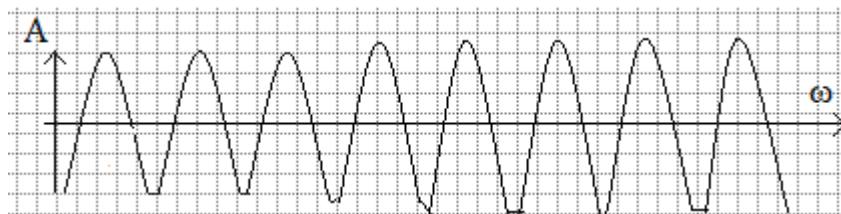


Рисунок 5 – График, иллюстрирующий работу высокочастотных шельфовых фильтров, в режиме усиления сигнала

- Режекторные (полосно-заграждающие) фильтры – удаляют частоты заданной полосы (рис. 6).

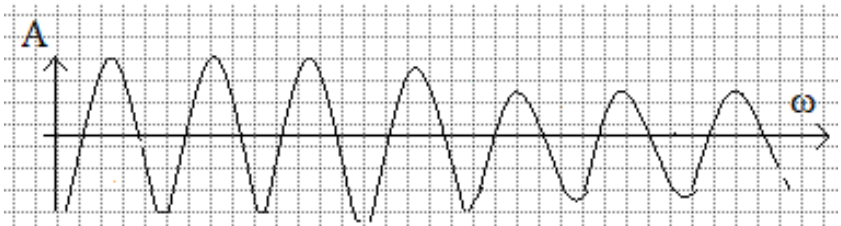


Рисунок 6 – График, иллюстрирующий работу режекторных фильтров

Для объективности необходимо рассмотреть типы линейных фильтров:

1) С конечной импульсной характеристикой – такие фильтры ограничены во времени по импульсной характеристике, не рекурсивны, так как в данных фильтрах выход не связан со входом, его АЧХ получается постоянной в полосе пропускания, в отличие от фазо-частотной характеристики (ФЧХ). Для устранения проблемы с ФЧХ используют групповые задержки [4]:

$$\frac{dW_0}{dW} = Tg,$$

где dW_0 – средний угол сдвига фазы,
 dW – средняя частота,
 Tg – групповая задержка.

Приведенная формула позволяет вывести необходимое условие для линейности ФЧХ сигнала, но использование дополнительных расчетов приводит к ещё большему временному ограничению импульсной характеристики.

2) С бесконечной импульсной характеристикой – в сравнении с предыдущим типом фильтра – данный, по импульсной характеристике, не ограничен во времени, обладает рекурсивностью, что обуславливает бесконечную длину во временной области на передаточной характеристике. Данная характеристика фильтров является более подходящей, при преобразовании звуков.

Реализовывать цифровые фильтры возможно на следующих уровнях:

Аппаратном – платформой является программируемая логическая интегральная схема (ПЛИС), преимущество заключается в быстродействии и легкости реализуемости в любом нужном виде [2].

Микропрограммном – предполагает расположение на процессорах и микроконтроллерах. Программа имеет преимущество в легкости воплощения изменений, настроек и т.п. задач, отвечающие за работу микроконтроллеров [2].

Из описанного сравнения получается, что для выбора подходящей комбинации фильтров, алгоритма и методов оцифровывания необходимо варьировать данные инструменты по отношению к ожидаемому результату и ресурсам вычислительных мощностей, так как каждый имеет свои преимущества в той или иной степени, влияющей на

качество обработанных данных. Для нахождения эффективной комбинации, необходимо изучить следующие методы:

- метод прямой оцифровки – преобразовывает сигнал, с помощью импульсно-кодовой модуляции. При этом преобразовании звук не изменяется по уровню звукового сигнала, а происходит нечто схожее на копирование звуковых импульсов. Но в результате, за счет роста объема данных – возникает снижение быстродействия обработки сигнала. Для решения данного недостатка используются алгоритмы сжатия, такие как MPEG, WMA и т.п.;

- командное кодирование – использование такого набора команд, который заменяет звучание, например музыкального инструмента, уровень, значение и продолжительность. Использование команд реализует генерирование данных в цифровой вид намного быстрее предыдущего. Стоит отметить, что для работы с преобразованным материалом (воспроизведение, запись) нужны специальные программы, которые выполняют одинаковые команды по-разному. Данный метод используется преимущественно для записанных звуков, в которых присутствуют шумы, помехи. Преимущество такого кодирования заключается в том, что количество задействованной памяти – невелико, что достигается возможностью использования одной команды неограниченное количество раз.

- комбинированный метод – использует преимущество прямой оцифровки с командным кодированием, путем сохранения семпла каждого звука, совместно с остальными командами. С учетом дополнения, памяти используется меньше, чем при первом способе. При этом разница звучания, из-за смены одной программы воспроизведения музыки, кодированной способом команд - отсутствует.

Алгоритмы оцифровывания сигнала

Для определения эффективного алгоритма оцифровывания сигнала, необходимо составить условие, к обрабатываемым данным, с учетом что каждый параметр влияет на качество выходного сигнала.

Прежде всего преобразованные данные необходимы для обработки речевых звуков. Первое условие это параметр дискретизации. Фундаментальное условие дискретизации было выведено в теореме Найквиста-Шеннона, которое звучит таким образом, что «Минимальная частота дискретизации должна быть min в два раза более частоты сигнала, а спектр должен быть не более заданного некоторого максимального значения частоты (частоты Найквиста)» для преобразований сигнала без искажений. Таким образом дискретизированный сигнал в 40 КГц имеет два оцифрованных звуковых фрагментов, при соотношении к аналоговому в 20 КГц. Для стандартизации максимальной частоты для систем аналогового цветного телевидения, во всем мире, принята величина в 44,1 КГц. Для DVD-дисков такой стандарт равен 48 КГц. В современной аудио-видео технике используется частота дискретизации от 96 КГц [1].

При несоблюдении условия минимальной частоты дискретизации, возникают искажения такого кодирования — Алиасинг (наложением спектра). К примеру, при частоте сигнала в 30 КГц частота дискретизации будет не 60 КГц, а 44 КГц – тогда выходной сигнал будет сильно искажен, что неприемлемо для нормального восприятия звука. Для нейтрализации таких негативных эффектов срезают частоты, выше 22 КГц, используя соответствующие фильтры.

После устранения отрицательно влияющих параметров на обработку сигнала, следует переходить к определению нужного алгоритма. Рассмотрение, стоит начинать с линейного способа квантования (импульсно-кодовой модуляции), что подразумевает округление каждого значения амплитуды до ближайшего, заданного по уровню, значения квантования. Точность округления зависит от выбранного заранее количества уровней и шага между ними. При использовании шестнадцатитбитного уровня квантования, погрешности становятся не заметны для слушателя

Использование неоднородной импульсно-кодовой модуляции, означает использование логарифмического метода разбиения амплитудной шкалы на различные фрагменты уровней. Отличие от предыдущего алгоритма заключается в неравномерном распределении уровней

квантования по всей амплитуде, что означает размещение большого количества уровней в области со низкой амплитудой.

Преимущество алгоритма состоит в уменьшении соотношения шума к сигналу для сигналов низкого уровня.

Адаптивная дифференциальная импульсно-кодовая модуляция, предполагает использование дифференциальных значений на всех этапах масштабирования квантования. Адаптируемость заключается в попытке предсказать следующее значение на основе разницы между реальным и предсказанным значениями. Использование размера уровня квантования вплоть до шестидесяти четырёх килобит повышает значительно скорость квантования.

Преобразование на аппаратном уровне происходит почти, как и в программной реализации, но с некоторой разницей. Первый этап состоит в соблюдении условия для дискретизации, который заключается в не превышении значения половины от её частоты. Это достигается с помощью фильтра низких частот. Затем выполняется замена непрерывного сигнала, его значениями в дискретные периоды времени, иными словами – дискретизация. Данное действие реализуется с помощью устройства выборки и хранения, которое располагается на выходе аналого-цифрового преобразователя. Стоит отметить, что разрядность этого устройства равняется двадцати четырём битам, исходя из эргономических требований современности. Затем, следует выполнить квантование сигнала. После чего на выходе получаем оцифрованный сигнал.

Выводы

Метод прямой оцифровки имеет высокое быстродействие за счет хранения каждой копии импульса, что требует много свободного места на запоминающем устройстве. Но метод командного кодирования позволяет сохранить свободное место, используя один сигнал необходимое количество раз, за счет чего теряется быстродействие, что не удовлетворяет условиям. Для решения недостатков стоит воспользоваться комбинированным методом, который при экономии свободного места сохраняет быстродействие, при обработке звука.

Наиболее эффективным алгоритмом можно считать адаптивную импульсно-кодовую модуляцию, с самой высокой скоростью квантования и максимальной точностью, в сравнении с другими способами.

При использовании комбинированного метода и линейного фильтра с бесконечной характеристикой, можно добиться максимального быстродействия при выполнении преобразования аналогового сигнала в цифровой. Если при этом использовать адаптивную импульсно-кодовую модуляцию, то качество оцифрованного звука станет наиболее высоким.

Исследование оцифровывания звуковых данных послужит основой для реализации приложения распознавания речи. Полученные результаты будут необходимы на этапе определения свойств цифрового сигнала. Также, подобранные, наиболее эффективные методики и алгоритмы, повысят быстродействие и качество обрабатываемых данных в приложении, в дальнейшем.

Литература

1. Рыбин С.В. Основы компьютерной обработки звука// Компьютерные инструменты в образовании: журн. 2000. URL: <http://cte.eltech.ru/ojs/index.php/kio/article/view/563> (дата обращения: 19.04.2021)
2. Абдуллаев У. М. Методы и алгоритмы обработки звуковых сигналов // Бюллетень науки и практики. 2020. Т. 6. №6. С. 25-30. <https://doi.org/10.33619/2414-2948/55/03>
3. Звукорежиссёр: журн. / Основы психоакустики, 2007. ч 11.
4. Фетюхина Л. В., Бутова О. А. Изучение методов анализа линейной дискретной системы: метод. ук. 2007. с17. URL: http://repository.kpi.kharkov.ua/bitstream/KhPI-Press/33283/1/prohramy_2017_Izuchenie_metodov_analiza.pdf (дата обращения: 20.04.2021)