

01, январь 2016

УДК 534.4

Программный синтез звука музыкальных инструментов

Тишина Е. М., студент

Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Информационные системы и телекоммуникации»

Хабрат М. Д., студент

Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Информационные системы и телекоммуникации»

Яковенко М. А., студент

Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Информационные системы и телекоммуникации»

Научный руководитель: Тихомирова Е. А., доцент
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Информационные системы и телекоммуникации»

elizarti@bmstu.ru

В статье описывается реализация аддитивного синтеза музыкального инструмента флейты.

Аддитивный синтез - это метод получения сложного тембра сложением различного количества простых волн. Аддитивный синтез часто называют синтезом Фурье. Французский математик Жозеф Фурье предложил идею получения волн сложной формы путем сложения простых синусоидальных волн. Пример реализации этого метода приведен на рис. 1.

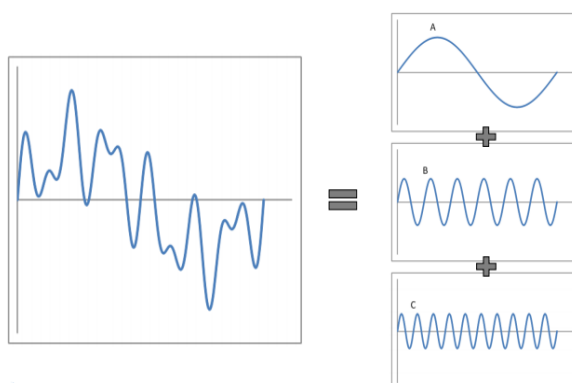


Рис. 1. Графическое изображение аддитивного синтеза

Для работы со звуковыми файлами был выбран формат WAV (Waveform Audio File Format). WAV — это формат для хранения несжатого аудиопотока, он широко используется в медиаиндустрии. Его особенность заключается в том, что для кодирования амплитуды волны выделяется фиксированное число бит. Это сказывается на размере выходного файла (он становится больше), но делает его очень удобным для чтения. Типичный WAV-файл представлен на рис. 2. Он состоит из трех частей: заголовка, тела с аудиопотоком (блок «data») и «хвоста» для дополнительной информации, куда можно вносить собственные метаданные.

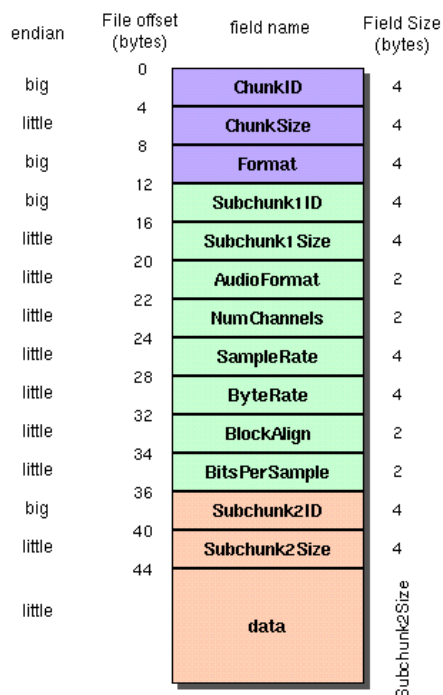


Рис. 2. Структура WAV файла

В заголовке содержатся основные параметры - число каналов, битрейт, число фреймов — на основании которых осуществляется разбор аудиопотока. WAV-файл содержит в себе 1 или 2 канала, каждый из которых кодируется 8, 16, 24 или 32 битами. Эти данные важно учитывать как при написании кода основной программы, так и при анализе звукового файла в программном пакете Adobe Audition.

В программе используется формат WAV как для анализа исходных звуковых файлов, так и для сохранения и воспроизведения синтезированных. В частности, для частотного анализа используется формат WAV PCM (Pulse Code Modulation), так как программно реализованный метод быстрого преобразования Фурье (БПФ) принимает на вход только PCM-файлы.

Для того чтобы передать данные на динамик, их следует сначала записать в WAV-файл. Для этого в потоке прописываются заголовки полей и информация о файле (например, частота дискретизации, битрейт и т. д.). Каждый элемент массива амплитуд разбивается на байты, переводится в аудиопоток и записывается в блок «data».

Главная сложность синтеза звука, состоит в том, что для исследуемого инструмента необходимо изучить модель звуковой волны. Для каждого музыкального инструмента модель индивидуальна и, чтобы синтезировать звук, нужно знать форму волны и тонкости её звучания. На рис. 3 приведены формы звуковых волн различных музыкальных инструментов.

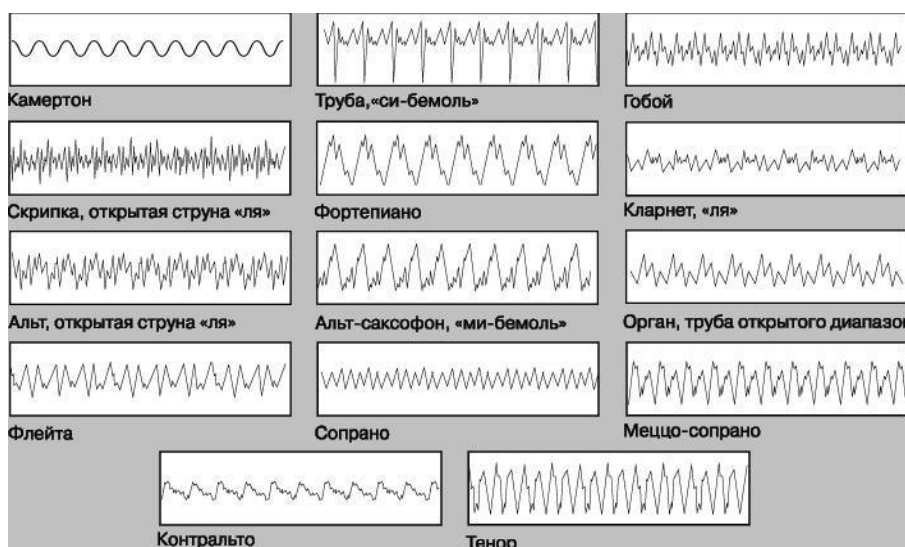


Рис. 3. Примеры форм звуковых волн различных музыкальных инструментов

Для того чтобы узнать математическое описание звуковой волны исследуемого инструмента можно использовать различные подходы. В данной статье для определения формы волны был использован фрагмент звучания классической флейты, которая имеет основной тон С4 (нота «Ля» 4 октавы). Фрагмент был найден в сети Интернет. Далее этот музыкальный отрывок был исследован с помощью программного пакета Adobe Audition. На рис. 4 продемонстрирована его амплитуда.

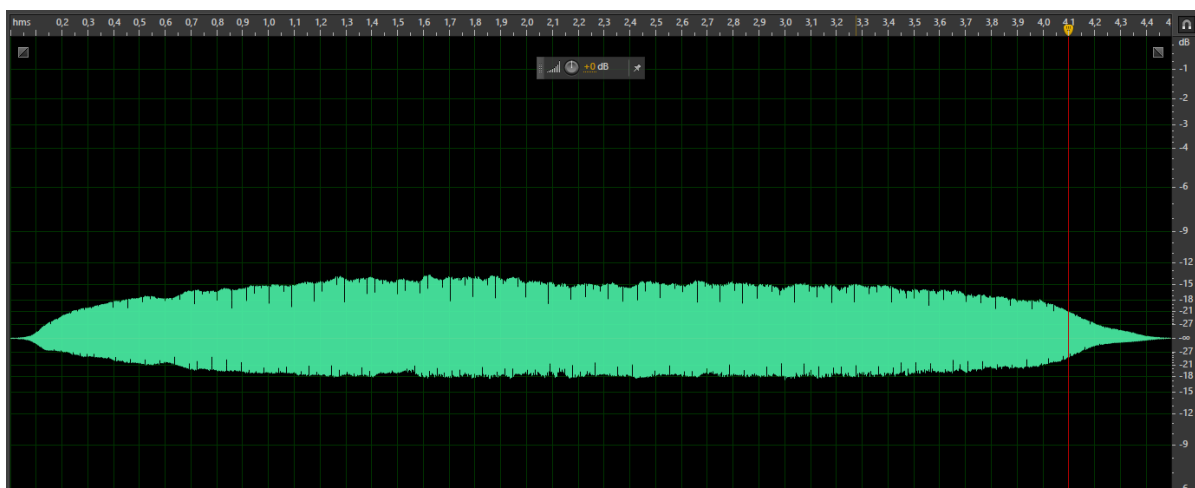


Рис. 4. Амплитуда звука, полученная в программном пакете Adobe Audition

На рис. 5 видны обертоны, которые возникают на частотах кратных основному тону. Для данного фрагмента основной тон составляет 260 Гц. Это число очень важно для дальнейшей реализации разрабатываемой программы.

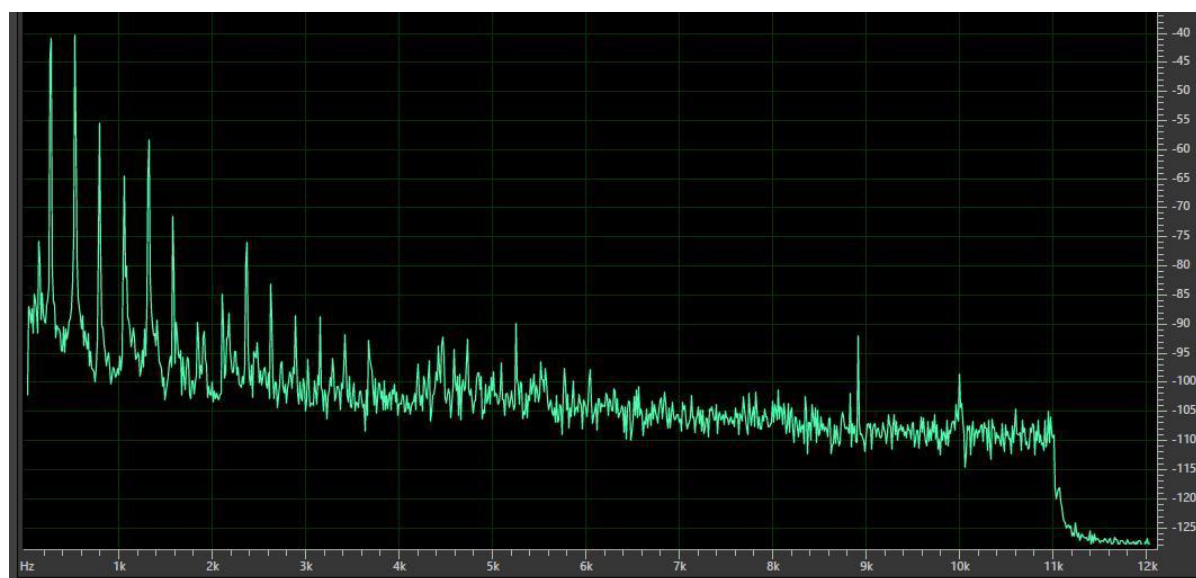


Рис. 5. Распределение обертонов по частотам, полученное в Adobe Audition

Для поиска амплитуд каждого обертона разработана вспомогательная программа. Для начала нужно выбрать достаточное количество отсчётов. Экспериментальным путём было установлено, что стоит рассматривать не менее 15 отсчётов. При меньшем количестве отсчётов интерполяция даёт слишком неточные результаты, то есть исходная функция слишком сильно отличается от интерполированной.

На заключительном этапе настройки вспомогательной программы указываются обертона, их может быть указано разное количество, но чем их больше, тем ближе синтезированный звук будет похож по звучанию на исходный. По определению значение частот гармонических обертонов кратно частоте основного тона, то есть в данном случае 260 Гц, и именно эти значения: 260 Гц, 520 Гц, 780 Гц и т.д. указываются на заключительном этапе настройки вспомогательной программы. В результате получается выходной файл, в котором можно увидеть амплитуды (их количество равно количеству отсчётов) для каждого обертона.

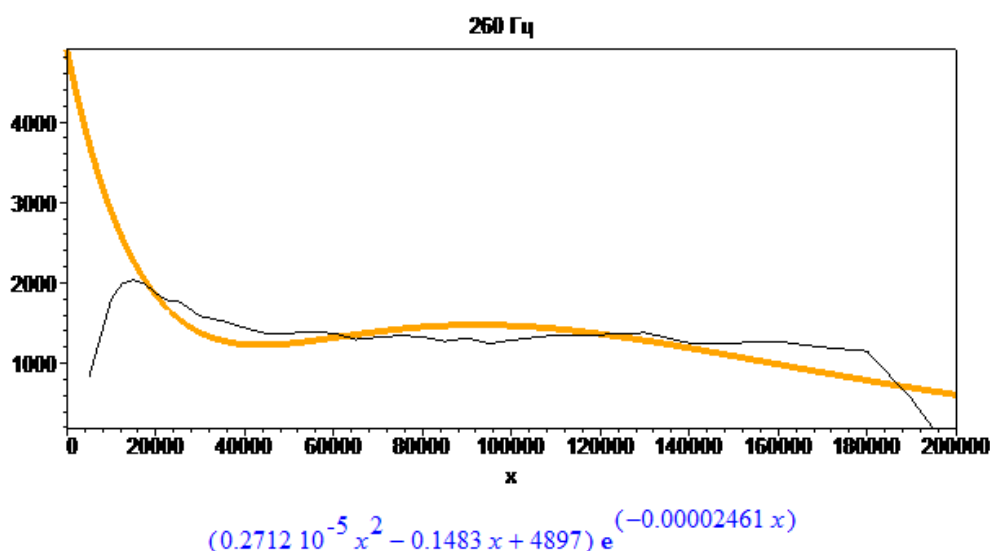
Теперь, когда статистика по амплитудам собрана, нужно найти функцию, которая бы смогла бы максимально точно описать все амплитуды по каждому обертому. Для обработки данных была выбрана система компьютерной алгебры Maple 17. Эксперименты с одной из функций для интерполяции NonlinearFit показали, что кривые неплохо приближаются следующим выражением:

$$y = (A \cdot x^2 + B \cdot x + C) \cdot e^{D \cdot x},$$

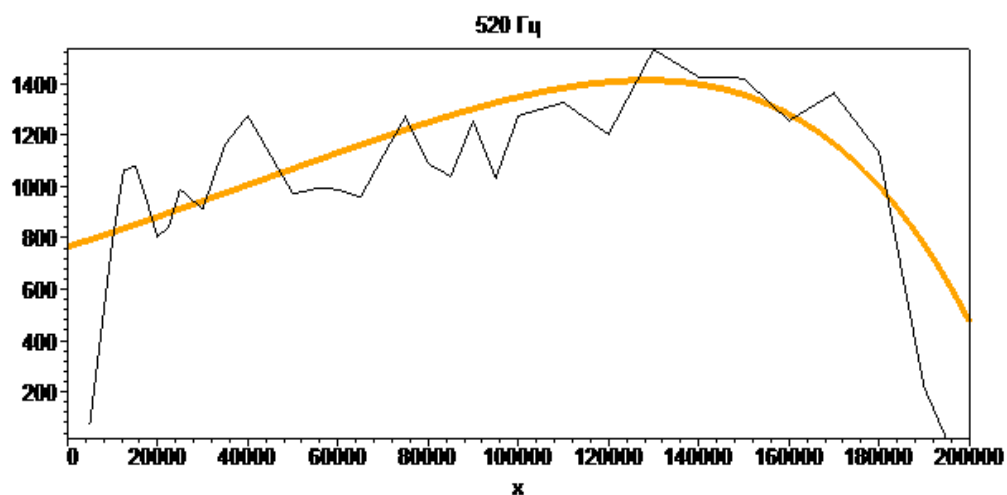
где A, B и C — некоторые коэффициенты многочлена, а D — коэффициент при степени экспоненты. Значения этих четырёх коэффициентов нужно найти для каждой исследуемой частоты для их последующей обработки в основной программе.

Повышение степени многочлена, а также попытки использовать периодические функции \sin или \cos не дали более точного результата при интерполяции.

Результаты интерполяции для первых гармоник приведены на рис. 6.

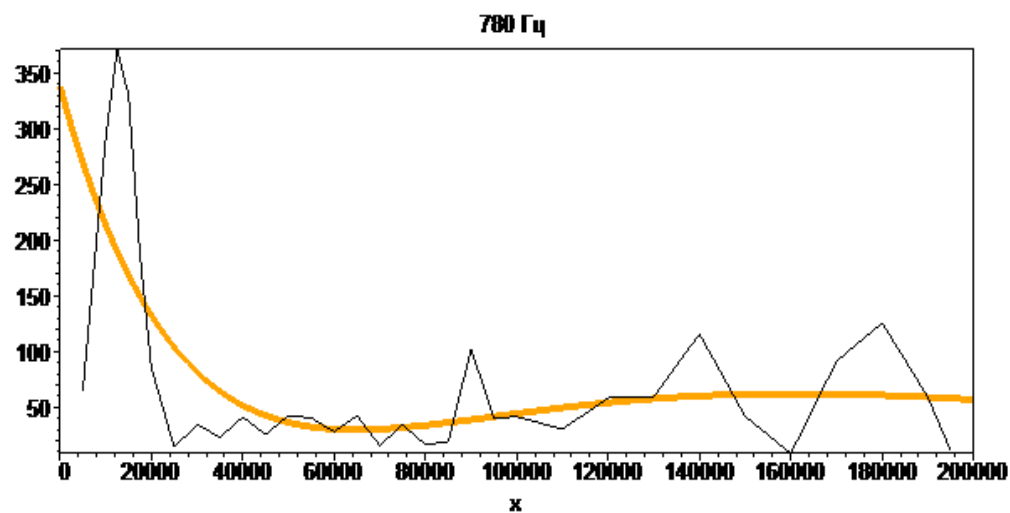


a)



$$(-0.4467 \cdot 10^{-8} x^2 - 0.002659 x + 734.2) e^{(0.00001089 x)}$$

б)



$$(0.8326 \cdot 10^{-7} x^2 - 0.009193 x + 337.8) e^{(-0.00001736 x)}$$

в)

Рис. 6. Интерполяция первых гармоник для различных значений частоты: а) значение частоты 260 Гц; б) значение частоты 520 Гц; в) значение частоты 780 Гц

Итак, функции, описывающие амплитуды, получены. Теперь эти данные можно внести в код основной программы. На рис. 7 приведён алгоритм работы программы.

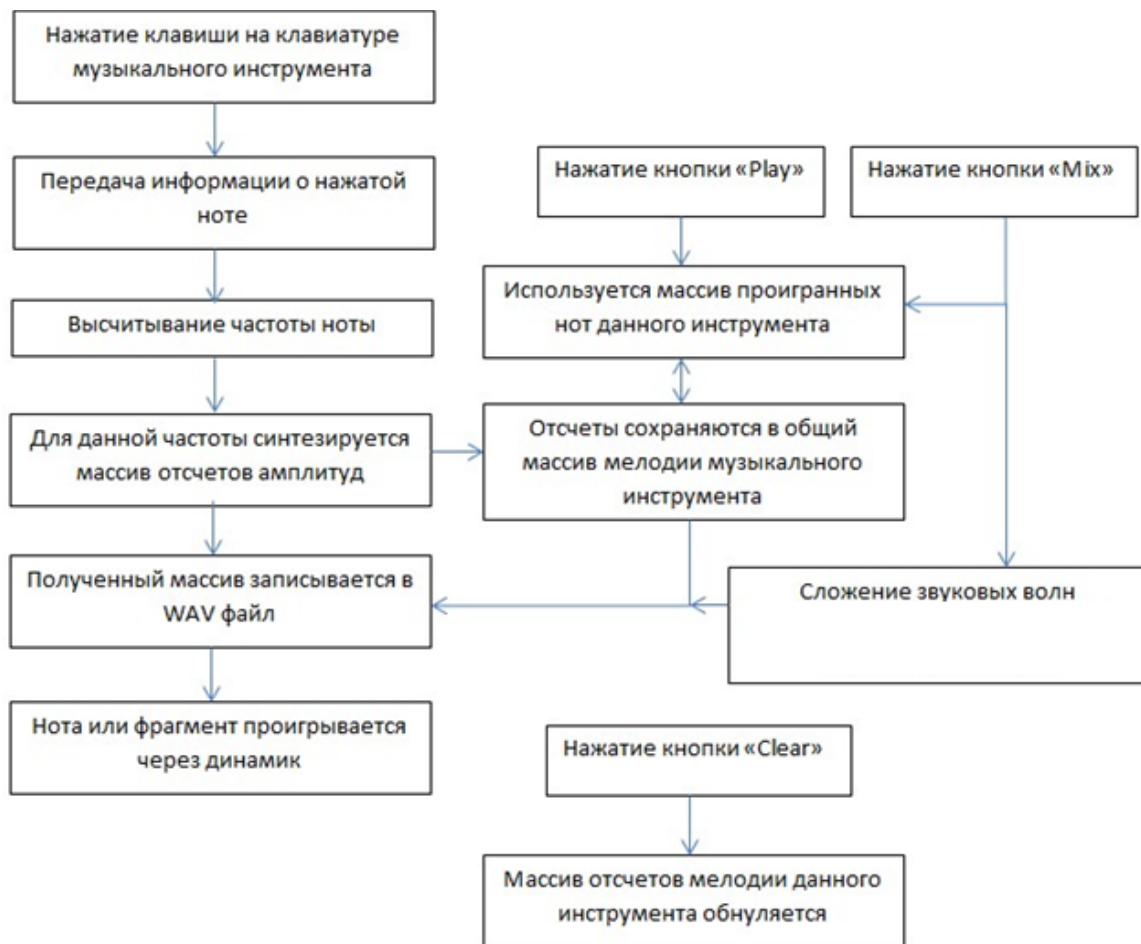


Рис. 7. Алгоритм работы программы

Для воспроизведения нот флейты разработан графический интерфейс, изображенный на рис. 8. В окно программы были помещены 2 октавы по 12 клавиш. При нажатии на одну из клавиш обработанный массив амплитуд помещается в блок «data» WAV-файла, после чего записанный WAV-файл воспроизводится.

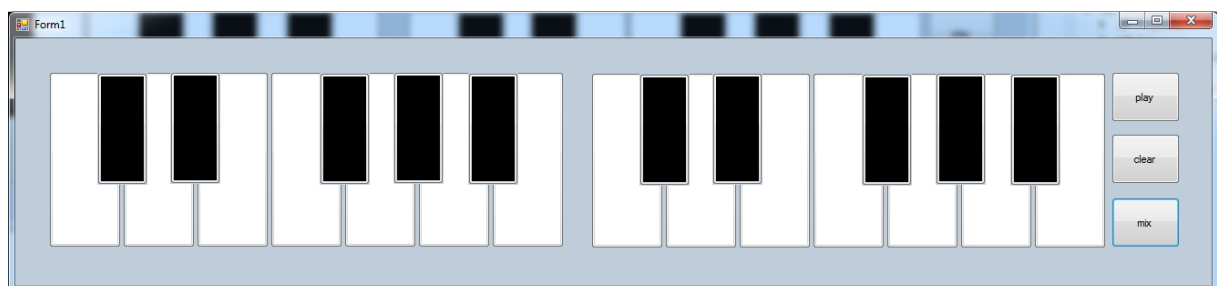


Рис. 8. Интерфейс разработанной программы

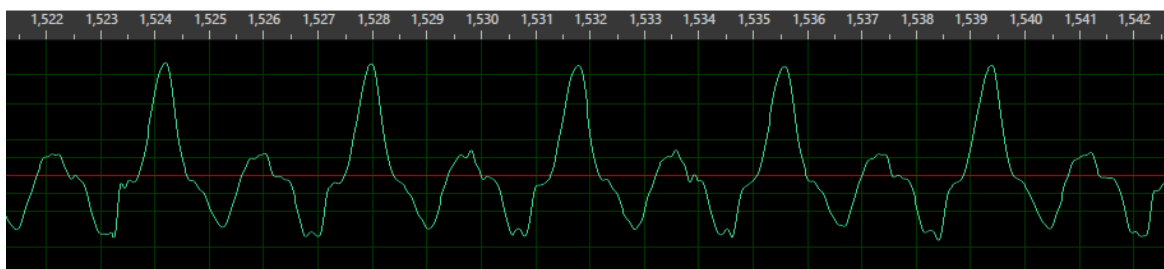


Рис. 9. Форма волны исходного звукового файла

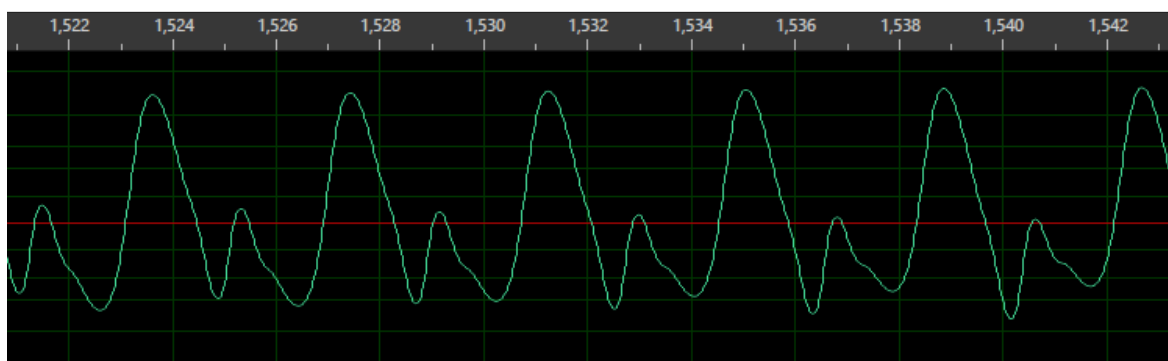


Рис. 10. Форма волны синтезированного звукового файла

Из рис. 9 и рис. 10 видно, что форма волны исходного звукового файла и форма волны синтезированного звукового файла очень похожи, но не идентичны друг другу. Эти различия вероятно вызваны недостаточным количеством выбранных отсчетов и погрешностью в процессе интерполяции.

Заключение

В статье была описана реализация программного синтеза музыкального инструмента флейта с помощью аддитивного синтеза. Полученный звуковой файл аудиально схож с первоначальным файлом. Однако визуально полученный график звуковой волны не идентичен исходному. Для получения наилучшего результата можно использовать другие способы интерполяции и увеличить количество анализируемых отсчетов.

Был реализован графический пользовательский интерфейс программы, который позволяет воспроизводить отдельные ноты, а также записывать, проигрывать и удалять уже записанные фрагменты мелодии. Осуществлена возможность воспроизведения отдельных нот на виртуальной музыкальной клавиатуре с физической клавиатуры.

Список литературы

- [1] Романова И.К. Анализ сигналов. Ряды Фурье и преобразование Фурье [Электрон. ресурс]: метод. указ. к лаб. работе по дисциплинам «Моделирование и исследование мехатронных систем», «Управление в технических системах». М.: МГТУ им. Н.Э. Баумана, 2014. 1 CD-ROM.
- [2] Герман Д.Я. Цифровая обработка сигналов [Электрон. ресурс]: учеб. пособие по дисциплине «Цифровая обработка сигналов». М.: МГТУ им. Н. Э. Баумана, 2010. 1 CD-ROM.
- [3] Галанин М.П., Савенков Е.Б. Методы численного анализа математических моделей. М.: МГТУ им. Н.Э. Баумана, 2010. 590 с.
- [4] Фурье-анализ и аддитивный синтез звука струны акустической гитары. Режим доступа: <https://blog.amartynov.ru/archives/fft-анализ-и-аддитивный-синтез-гитары/> (дата обращения 20.05.15).