**ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ**

**НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ**

**«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

Факультет компьютерных наук

Образовательная программа бакалавриата «Программная инженерия»

|  |  |  |
| --- | --- | --- |
| СОГЛАСОВАНО  Руководитель проекта,  профессор департамента больших  данных и информационного поиска  ФКН, доктор физ.-мат. наук  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ А. Б. Шаповал  «\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_ 2021 г. |  | УТВЕРЖДАЮ  Академический руководитель образовательной программы «Программная инженерия»  профессор департамента программной инженерии, канд. техн. наук  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ В.В. Шилов  «\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_ 2021 г. |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | ***Подп. и дата*** |  | | ***Инв. № дубл.*** |  | | ***Взам. инв. №*** |  | | ***Подп. и дата*** |  | | ***Инв. № подл*** |  | | **Аудиоплагин для создания стереозвука**  **Пояснительная записка**  **ЛИСТ УТВЕРЖДЕНИЯ**  **RU.17701729.04.01-01 81 01-1-ЛУ** | | |
|  |  | |
| Исполнитель  студент группы БПИ 196  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ /М.С. Шестаков /  «\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2021 г. | |
|  | | |
|  | |  |

**Москва 2021**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| УТВЕРЖДЕН  RU.17701729.04.01-01 81 01-1-ЛУ |  | |  | |
| |  |  | | --- | --- | | ***Подп. и дата*** |  | | ***Инв. № дубл.*** |  | | ***Взам. инв. №*** |  | | ***Подп. и дата*** |  | | ***Инв. № подл*** |  | | **Аудиоплагин для создания стереозвука**  **Пояснительная записка**  **RU.17701729.04.01-01 81 01-1**  **Листов ??** | | | | |
|  | |  | | |
|  | | |
|  | | | | |
|  | | | |  |

**Москва 2021**

**СОДЕРЖАНИЕ**

Оглавление

[1) ВВЕДЕНИЕ 4](#_Toc72137273)

[1.1 Наименование программы 4](#_Toc72137274)

[1.2 Документы, на основании которых ведется разработка программы 4](#_Toc72137275)

[2. НАЗНАЧЕНИЕ И ОБЛАСТЬ ПРИМЕНЕНИЯ 5](#_Toc72137276)

[2.1 Назначение программы 5](#_Toc72137277)

[2.1.1 Функциональное назначение 5](#_Toc72137278)

[2.1.2 Эксплуатационное назначение 5](#_Toc72137279)

[2.2 Краткая характеристика области применения 5](#_Toc72137280)

[3. ТЕХНИЧЕСКИЕ ХАРАКТЕРИСТИКИ 6](#_Toc72137281)

[3.1 Постановка задачи на разработку программы 6](#_Toc72137282)

[3.2 Применяемые математические методы в программе 6](#_Toc72137283)

[3.2.1 Оконное преобразование Фурье 6](#_Toc72137284)

[3.2.2 Получение фазы и амплитуды 6](#_Toc72137285)

[3.2.3 Обратное преобразование 6](#_Toc72137286)

[3.3 Описание и обоснование алгоритма и функционирования программы 7](#_Toc72137287)

[3.3.1 Накопление блоков данных 7](#_Toc72137288)

[3.3.2 Применение преобразования Фурье 7](#_Toc72137289)

[3.3.3 Выделение амплитуды и фазы для каждой из частот 8](#_Toc72137290)

[3.3.4 Разделение частот на стерео 8](#_Toc72137291)

[3.3.5 Преобразование амплитуды и фазы обратно в комплексное число 9](#_Toc72137292)

[3.3.6 Обратное преобразование Фурье 9](#_Toc72137293)

[3.3.7 Применение WOLA 9](#_Toc72137294)

[3.3.8 Накопление полученных данных и возврат результата в DAW 9](#_Toc72137295)

[3.4 Описание и обоснование выбора метода организации входных и выходных данных 9](#_Toc72137296)

[3.4.1 Описание метода организации входных и выходных данных 9](#_Toc72137297)

[3.4.2 Обоснование метода организации входных и выходных данных 10](#_Toc72137298)

[3.5 Описание и обоснование выбора состава технических и программных средств 10](#_Toc72137299)

[4. ТЕХНИКО-ЭКОНОМИЧЕСКИЕ ПОКАЗАТЕЛИ 11](#_Toc72137300)

[4.1 Предполагаемая потребность 11](#_Toc72137301)

[4.2 Экономические преимущества разработки по сравнению с отечественными и зарубежными образцами или аналогами 11](#_Toc72137302)

[ПРИЛОЖЕНИЕ 1 ТЕРМИНОЛОГИЯ 12](#_Toc72137303)

[ПРИЛОЖЕНИЕ 2 13](#_Toc72137304)

[ПРИЛОЖЕНИЕ 3 14](#_Toc72137305)

[СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ 14](#_Toc72137306)

[ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ 16](#_Toc72137307)

# ВВЕДЕНИЕ

## Наименование программы

Наименование программы ̶ "Аудиоплагин для создания стереозвука"

Наименование программы на английском языке ̶ "Audioplugin for creating stereo"

Краткое наименование программы – "Стерео-плагин" ("Stereo Plugin”)

## Документы, на основании которых ведется разработка программы

Разработка программы ведётся на основании приказа декана факультета компьютерных наук И.В. Аржанцева «Об утверждении тем, руководителей курсовых работ студентов образовательной программы «Программная инженерия» факультета компьютерных наук» № 2.3-02/1112-04 от 11.12.2020.

# НАЗНАЧЕНИЕ И ОБЛАСТЬ ПРИМЕНЕНИЯ

## Назначение программы

### Функциональное назначение

Программа предназначена для преобразования моно звука в стерео без значительных потерь качества при обратном преобразовании.

### Эксплуатационное назначение

Программа представляет собой плагин для цифровой звуковой рабочей станции (DAW), предназначенный для использования создателями электронной музыки.

## Краткая характеристика области применения

В современном мире большой процент музыки люди слушают в наушниках. А поскольку наушники способны выдавать два отдельных звуковых потока (стерео звук), то современным композиторам нужно создавать музыку с двумя аудиодорожками. Соответственно, возникает задача как сделать из одной аудиодорожке две так, чтобы они звучали как можно более объёмно. Разрабатываемая программа призвана решить данную проблему.

При этом важно понимать, что при прослушивании музыки не в наушниках, звуковые потоки из правой и левой дорожки будут попадать в оба уха, то есть, по сути, происходит обратное преобразование стерео звука в моно. Поэтому разрабатываемая программа должна стремиться к тому, чтобы получающийся в результате обратного преобразования звук был максимально близок к исходному.

# ТЕХНИЧЕСКИЕ ХАРАКТЕРИСТИКИ

## Постановка задачи на разработку программы

Программа должна принимать из DAW блоки данных со звуком (посредством технологии VST), с помощью математических методов преобразовывать их в стерео (то есть разбивать на две звуковые дорожки), и возвращать обратно в DAW. В программе должна быть предусмотрена возможность менять параметры преобразования посредством пользовательского интерфейса.

## Применяемые математические методы в программе

### Оконное преобразование Фурье

Для преобразования звука из временного пространства в частотное в программе используются оконные преобразования Фурье. Для расчёта единоразового преобразования Фурье в момент времени t для частот используется следующая формула:

В данной формуле – значение звукового сигнала в точке n (звуковой сигнал центрирован относительно момента времени t, N – количество значений сигнала, обрабатываемых за одно преобразование, w[n] – некоторая весовая (оконная) функция, R – сдвиг между соседними преобразованиями.

Для оконной функции выбрано прямоугольное окно в связи с его максимально широким спектром принимаемых частот и меньшими потерями при внесении изменений в .

Для количества значений, обрабатываемых за единичное преобразование, было взято значение N = 2048. В качестве сдвига между соседними преобразованиями Фурье было выбрано число R=1024, таким образом достигается перекрытие в 50%.

### Получение фазы и амплитуды

Пусть имеется некоторая частота . Тогда соответствующее ей значение будет равно:

Где N – количество значений сигналов, обрабатываемых за одно преобразование Фурье, а – количество сэмплов звуковой волны, проигрываемых в секунду. Данная формула следует из частоты Найквиста.

Амплитуда для данной частоты в момент времени t будет равна:

Фаза, в свою очередь, равна:

### Обратное преобразование

Для обратного преобразования используется следующая формула:

Где нормирующая константа K в общем случае вычисляется как , но для прямоугольного окна получается равна 1.

Затем полученные смешиваются в с помощью метода weight-overlap-add (WOLA).

В качестве весовой функции для WOLA была взята функция Хеннинга:

При пересечении в 50% соседние функции в сумме дают 1 и не нуждаются в дополнительной нормировке.

## Описание и обоснование алгоритма и функционирования программы

Процесс обработки данных в программе состоит из нескольких этапов:

* 1. Накопление блоков данных
  2. Применение преобразования Фурье
  3. Выделение амплитуды и фазы для каждой из частот
  4. Разделение частот на стерео
* Генерация маски
* Применение маски к амплитудам
  1. Преобразование амплитуды и фазы обратно в комплексное число
  2. Обратное преобразование Фурье
  3. Применение WOLA
  4. Накопление полученных данных и возврат результата в DAW

### Накопление блоков данных

Согласно спецификации формата VST, входные данные поступают в аудиоплагин блоками переменной длины. Поскольку для применения преобразований Фурье необходимы блоки одинаковой длины, то данные необходимо накапливать. Это поведение реализовано с помощью класса DataAccumulator.

Сложность данного решения по времени – O(N), по памяти O(1), где N – количество поступающих сэмплов.

### Применение преобразования Фурье

В силу особенностей работы человеческого слуха человек слышит не всю волну целиком, а воспринимает отдельные частоты. Поэтому для создания стерео звука также необходимо представить звук в виде частот. Для этого были выбраны преобразования Фурье, которые обладают оптимальной скоростью и качеством, а также очень распространены при анализе звука.

Алгоритмическая сложность для быстрых преобразований Фурье – O(K log K), где K – выбранная длина для преобразований Фурье.

Алгоритмическая сложность на промежутке в N сэмплов долей пересечения x – O(N / x). (K в программе было выбрано константой)

### Выделение амплитуды и фазы для каждой из частот

В результате преобразований Фурье получаются некоторые комплексные числа, и для более удобной работы с ними необходимо выделить для каждой частоты амплитуду и фазу. Эти данные выделяются с помощью вычисления модуля и аргумента комплексного числа. Также необходимо по частоте в герцах определять соответствующую ей амплитуду в массиве амплитуд. За это отвечает функция *MainProcessor::FFTIndexToFreq*.

### Разделение частот на стерео

Наложенное на программу ограничение – сохранение обратного преобразования в моно и отсутствие каких-либо дополнительных частот в стерео – оставляет лишь единственный способ преобразования звука в стерео: направить некоторые частоты из исходного звука в одну звуковую дорожку, а остальные в другую. При этом допускается смешанный вариант, когда часть частоты идёт в левую звуковую дорожку, а оставшаяся часть в правую.

Для удобства данное преобразование было разделено на два этапа:

1. Генерация маски

Для более равномерного звучания частоты распределяются по функции синус. При этом в программе существует два режима разделения:

* Динамический

В этом случае распределение частот динамически меняется с течением времени в зависимости от распределения амплитуд частот. При этом синус быстрее колеблется в тех точках, где амплитуда больше, и медленнее, где амплитуда меньше. В результате достигается баланс громкости между результирующими звуковыми дорожками.

При этом для плавности используется не мгновенная амплитуда, а амплитуды сглаженная. Для этого высчитывается массив средних амплитуд по формуле:

*mediumMagnitude[i] = weight \* mediumMagnitude[i] + (1 – weight) \* magnitude[i]*, где:

*magnitude[i]* – текущая частота,

*weight* – параметр сглаживания, задаваемый пользователем с помощью параметра “*Attack*”.

* Статический

В этом режиме распределение частот статично и не меняется с течением времени. Поскольку в музыке для частот принята логарифмическая шкала (то есть нота в каждой следующей октаве имеет частоту в два раза больше предыдущей), то распределение также

В обоих режимах пользователь может регулировать быстроту колебаний синуса с помощью регулятора “*Frequency Spread*”.

Выбор режима разделения определяется пользователем с помощью кнопки “*Dynamic Split*”.

Также для создания более естественного разделения, в программе есть возможность добавить в маску случайный шум. Чтобы шум был более плавный при генерации очередного значения для частоты x в временной промежуток t, его значение смешивается предыдущими (x-1, t) и (x, t-1). Силу применения шума к маске определяет пользователь с помощью регулятора “*Noise*”.

Генерация маски реализована посредством функции *MainProcessor::generateMask*.

1. Применение маски к амплитудам

Пусть имеется маска mask со значениями от 0 до 1 и исходные частоты data. Тогда разделение будет посчитано по следующей формуле:

*res1[i] = {data[i].magnitude \* mask[i], data[i].phase};*

*res2[i] = {data[i].magnitude \* (1 - mask[i]), data[i].phase};*

Применение маски реализовано посредством функции *MainProcessor::processSplit*.

### Преобразование амплитуды и фазы обратно в комплексное число

Для данного преобразования используется функция из стандартной библиотеки C++ -- *std::polar*.

### Обратное преобразование Фурье

Обратное преобразование Фурье имеет параметры и сложность, аналогичные прямому преобразованию Фурье. При этом данные, полученные после данного преобразования накапливаются в промежуточном буфере *MainProcessor::* *fftOutAccumulators,* который реализован с помощью класса DataAccumulator.

### Применение WOLA

Для вычисления результата необходимо скомпоновать данные, полученные в результате многочисленных преобразований Фурье. При этом данные на краях каждого из преобразований получаются менее точными, чем в середине. Поэтому для смешивания используется весовая функция Хэмминга. Само смешивание происходит в методе *MainProcessor::process*.

### Накопление полученных данных и возврат результата в DAW

Данные в WOLA также генерируются блоками постоянного размера, в то время как DAW может запрашивать блоки разного размера. Поэтому получаемые данные накапливаются в специальный *DataAccumulator*, данные из которого передаются в DAW. В результате работы плагина формируется некоторая задержка (латентность) между входными и выходными данными, равная длине преобразований Фурье (2048). Эта задержка сообщается в DAW с помощью функции *getLatencyInSamples*.

## Описание и обоснование выбора метода организации входных и выходных данных

### Описание метода организации входных и выходных данных

Входными данным для программы являются:

1. Входные данные, передаваемые из DAW посредством технологии VST.

- входящий аудиопоток (в формате моно или стерео).

- значения параметров плагина.

1. Действия пользователя.

- взаимодействие с пользовательским интерфейсом.

Выходными данными программы являются:

1. Выходные данные, передаваемые в DAW посредством технологии VST:

- результирующий аудиопоток (в формате стерео)

- базовая информация о плагине (название плагина, параметры и другие значения, которые требуется передавать согласно формату VST)

- значения параметров плагина

1. Выходные данные для пользователя.

- изображение пользовательского интерфейса

### Обоснование метода организации входных и выходных данных

Метод организации входных и выходных данных определила спецификация формата VST.

## Описание и обоснование выбора состава технических и программных средств

Для корректной и стабильной работы, программе необходимы следующие технические и программные средства:

1. Графическая карта: с поддержкой OpenGL 3.0;
2. Оперативная память: не менее 200 Мб свободной памяти;
3. Постоянная память: не менее 50 Мб свободной памяти на используемом накопителе;
4. Периферийные устройства: Клавиатура, мышь, а также устройство, способное выводить две звуковые дорожки раздельно (лучше всего наушники);
5. Аудиокарта;
6. USB;

Также должны быть выполнены все требования, необходимые для используемой совместно с программой DAW.

# ТЕХНИКО-ЭКОНОМИЧЕСКИЕ ПОКАЗАТЕЛИ

## Предполагаемая потребность

В современном мире огромный процент музыки слушается через наушники. Как следствие, при создании электронной музыке нужно каким-то образом разделять звук на два аудиопотока. Поэтому данный плагин может быть полезен практически любому композитору.

## Экономические преимущества разработки по сравнению с отечественными и зарубежными образцами или аналогами

Одним из основных методов для придания звуку объёмности является вставка задержки между левой и правой аудиодорожкой. Данный метод работает за счёт того что в реальной жизни звук тоже приходит в одно ухо немножко раньше, чем в левое, то есть между тем как мы услышим звук одним ухом и другим ухом есть некоторая задержка. Но у этого метода есть два серьёзных недостатка:

1. На самом деле в реальном мире звук отражается от стен, причём волны разных частот и направлений распространяются в воздухе по-разному. Поэтому на самом деле звук в левом ухе и правом ухе отличается не только задержкой. Из-за этого при использовании данного метода человек всё ещё может ощущать некоторую искусственность и нереалистичность.
2. Данный метод сильно искажает музыку, если преобразовать её обратно в моно звук. Задержка создаёт сильный эффект эхо, как будто бы музыку записывали в тесном помещении.

Именно этот метод используется в большинстве аналогов.

Создаваемая программа предполагает использование другого метода для преобразования, поэтому она не должна иметь данных недостатков.

Также стоит отметить, что данная программа будет распространяться бесплатно, в отличие от большинства аудиоплагинов.

# ПРИЛОЖЕНИЕ 1 ТЕРМИНОЛОГИЯ

# ПРИЛОЖЕНИЕ 2

|  |  |
| --- | --- |
| Классы | Назначение |
| AIController | Класс для игрока, управлемого искуственным интеллектом |

Таблица 3.1

Описание полей, методов и свойств класса AIController

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Методы класса AIController | | | | |
| Имя | Модификатор доступа | Тип | Аргументы | Назначение |
| Update | private | void |  | Изменяет параметры игрока на основе ИИ. Автоматически вызывается Unity каждый кадр |

# ПРИЛОЖЕНИЕ 3

# СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

1. ГОСТ 19.101-77 Виды программ и программных документов. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
2. ГОСТ 19.102-77 Стадии разработки. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
3. ГОСТ 19.103-77 Обозначения программ и программных документов. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
4. ГОСТ 19.104-78 Основные надписи. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
5. ГОСТ 19.105-78 Общие требования к программным документам. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
6. ГОСТ 19.106-78 Требования к программным документам, выполненным печатным способом. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
7. ГОСТ 19.201-78 Техническое задание. Требования к содержанию и оформлению. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
8. ГОСТ 19.603-78 Общие правила внесения изменений. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
9. ГОСТ 19.604-78 Правила внесения изменений в программные документы, выполненные печатным способом. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
10. ГОСТ 15150-69 Машины, приборы и другие технические изделия. Исполнения для различных климатических районов. Категории, условия эксплуатации, хранения и транспортирования в части воздействия климатических факторов внешней среды. – М.: Изд-во стандартов, 1997.
11. ГОСТ 19.602-78 Правила дублирования, учета и хранения программных документов, выполненных печатным способом. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
12. ГОСТ 19.301-79 Программа и методика испытаний. Требования к содержанию и оформлению. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.

[https://ccrma.stanford.edu/~jos/sasp/Overlap\_Add\_OLA\_STFT\_Processing.html]

[https://www.mathworks.com/help/signal/ref/istft.html]

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Лист регистрации изменений | | | | | | | | | |
| Номера листов (страниц) | | | | | Всего листов (страниц в докум.) | № документа | Входящий  № сопроводит ельного докум. и дата | Подп. | Дата |
| Изм. | Изменен ных | Заменен ных | Новых | Аннули рованных |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

# ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ