

АНАЛИЗ VST REAL-TIME ПЛАГИНОВ ДЛЯ ГЕНДЕРНОГО МОРФИНГА ГОЛОСА

Каширин Максим Александрович

Аспирант, Омский государственный
университет им. Ф.М. Достоевского;
Ведущий разработчик, ООО «Симбирсофт»
kashirin-maxim1998@yandex.ru

ANALYSIS OF VST REAL-TIME PLUGINS FOR VOICE GENDER MORPHING

M. Kashirin

Summary. The article discusses methods of changing sound using VST instruments in real time, with an emphasis on converting voices from male to female and vice versa. Popular tools such as MorphVOX Pro and Clownfish Voice Changer have been analyzed, which allow users to instantly change their voice in contexts such as gaming, streaming and voice chats. Special attention is paid to the fundamental frequencies and frequency ranges important for determining the differences between male and female voices, as well as how they can be modified using technologies such as pitch change and formant shift to achieve accurate and realistic sound conversion. Technical difficulties related to audio editing are also discussed, such as maintaining sound quality and realism during conversion, as well as distortion or delay issues when using real-time tools. These tools offer huge potential to enhance creativity in areas such as filmmaking and gaming, however, there is an ongoing need to develop more accurate and faster methods to overcome these challenges. In conclusion, the article highlights the importance of this technology for creating an innovative audio experience and points to areas of future research aimed at improving tools for changing sound using artificial intelligence, improving the quality of modified sound and reducing time delays.

Keywords: voice morphing, voice changing, vst plugins, digital audio processing, sounddesign.

Аннотация. В статье рассматриваются методы изменения звука с использованием VST (Virtual Studio Technology) инструментов в реальном времени, с акцентом на преобразование голосов с мужского на женский и наоборот. Проанализированы популярные инструменты, такие как MorphVOX Pro и Clownfish Voice Changer, которые позволяют пользователям мгновенно изменять голос в таких контекстах, как игры, потоковое вещание и голосовые чаты. Особое внимание уделяется основным частотам и диапазонам частот, важным для определения различий между мужскими и женскими голосами, а также тому, как их можно модифицировать с помощью таких технологий, как изменение высоты тона и формантный сдвиг для достижения точного и реалистичного преобразования звука. Также обсуждаются технические трудности, связанные с редактированием звука, такие как сохранение качества звука и реализма во время преобразования, а также проблемы искажения или задержки при использовании инструментов в реальном времени. Эти инструменты открывают огромный потенциал для повышения креативности в таких областях, как кинопроизводство и игры, однако существует постоянная необходимость в разработке более точных и быстрых методов для преодоления этих проблем. В заключение, статья подчеркивает важность этой технологии для создания инновационного аудиоопыта и указывает на области будущих исследований, направленных на улучшение инструментов для изменения звука с помощью искусственного интеллекта, повышение качества изменённого звука и сокращение временных задержек.

Ключевые слова: морфинг голоса, изменение голоса, vst плагины, цифровая обработка звука, саунддизайн.

Введение

В современную цифровую эпоху технологии стали неотъемлемой частью повседневной жизни, играя ключевую роль в совершенствовании многих сфер, включая медиа и индустрию развлечений. Одной из таких технологий является изменение голоса, которое активно применяется в кинопроизводстве, прямых трансляциях и видеоиграх. Эта технология позволяет пользователям не только изменять голос, но и полностью менять его идентичность. Одним из наиболее распространённых инструментов для изменения голоса является VST (Virtual Studio Technology) — программная платформа, обеспечивающая возможность добавления эффектов и модификаций к аудиозаписям в режиме реального времени. За последние годы технологии VST значительно эволюционировали, став более точными и ре-

алистичными, что позволяет как профессионалам, так и новичкам добиваться впечатляющих результатов [1].

VST-платформа представляет собой ключевой компонент программного обеспечения для записи и редактирования звука, предлагая широкий спектр инструментов для изменения звуковых характеристик. Среди наиболее популярных — плагины, используемые для изменения высоты голоса, модуляции частот (frequency shifting) или добавления таких эффектов, как эхо или искажение, что позволяет полностью изменить идентичность голоса.

Особую ценность технологии VST представляют для изменения голоса с мужского на женский и наоборот, что делает их полезными в различных креативных сферах. Например, в игровой индустрии такие технологии применяются для синхронизации голоса персонажа

с его ролью. В кино и сериалах они позволяют актёрам менять голос так, чтобы он звучал, как голос другого человека или противоположного пола [2].

Кроме того, в прямых трансляциях и подкастах ведущие и исполнители часто используют VST для изменения своего голоса в реальном времени, чтобы привлечь внимание аудитории или создать определённую атмосферу. В данной работе будут всесторонне рассмотрены возможности VST-инструментов для изменения голоса с мужского на женский и обратно, а также будут исследованы технические различия между голосами по тональности, основным частотам и другим акустическим аспектам. Также будут кратко описаны наиболее популярные программы, основанные на VST, и их практическое применение в различных областях, включая музыку, игры и кино

Цели

Данный проект ставит перед собой несколько основных целей, которые описаны ниже и связаны с использованием инструментов VST для изменения звука. Данная статья будет стремиться к следующему:

1. Понимание VST-инструментов, использующихся для редактирования звука. Данный пункт предполагает проведение обзора самых популярных инструментов VST, которые представлены на рынке, и являются основными инструментами в реальном времени, использующиеся для смены голоса. Акцент будет сделан на способах реализации этих инструментов, на их точность, удобство использования и гибкость выполнения для новичков и профессионалов;
2. Применение VST-технологий для работы в реальном времени. Проведение практической работы, соответствующей упомянутым инструментам, позволит увидеть применение теоретических данных на практике. В данном пункте будут описаны инструменты высотного и формантного сдвигов тона Change pitch и Formant changer.

Данные инструменты представляют собой смену тона голоса и изменение формантных частот, чтобы измененные изменения частот ассоциировались мужчинами и женщинами; С этими целями статья стремится предоставить всесторонний обзор того, как эффективно использовать технологию VST для изменения голоса, выделяя технические различия между мужскими и женскими голосами, а также предлагая практические решения для преодоления трудностей, с которыми сталкиваются пользователи при работе с этими инструментами.

Объяснение VST и его плагинов для работы в реальном времени

Это стандарт программного обеспечения, разработанный компанией Steinberg, который позволяет поль-

зователям интегрировать звуковые инструменты и эффекты в программы для редактирования цифрового аудио. Основу технологии составляют плагины, которые предназначены для обработки звука или генерации звуковых эффектов в цифровой среде. VST предоставляет пользователям возможность создавать сложные аудиозаписи, применять разнообразные эффекты и изменять звук в режиме реального времени. Это делает VST незаменимым инструментом в современном музыкальном производстве и аудиоредактировании, обеспечивая высокий уровень гибкости и качества звука [1-4].

Методы и материалы

Особенности VST [1–4]:

1. VST — универсальный инструмент, который находит применение в различных областях аудио, начиная от создания музыки и заканчивая звуковым монтажом для кино и трансляций. Возможности VST охватывают широкий спектр функций для редактирования звука, включая изменение тональности, настройку частот, добавление звуковых эффектов и полное преобразование звучания до неузнаваемости.
2. Тщательное управление звуком. Инструменты VST обеспечивают точный контроль над всеми аспектами звука. Например, пользователи могут регулировать основные частоты и тембры, что позволяет детализировать звучание для достижения желаемого результата как для голоса, так и для музыкальных инструментов.
3. Расширяемость. Благодаря множеству доступных плагинов VST предлагает гибкость в настройках. Плагины могут быть разработаны как сторонними компаниями, так и созданы самостоятельно, что делает VST очень адаптивным инструментом, позволяя пользователям изменять и расширять его функционал в соответствии с индивидуальными потребностями.

Функции плагинов VST для работы в реальном времени

Одной из ключевых особенностей VST инструментов является способность обрабатывать и изменять звук в реальном времени, что позволяет пользователям немедленно услышать изменения после их внесения, без необходимости ожидать обработки звука. Вот основные преимущества работы в режиме реального времени:

1. Моментальное и интерактивное редактирование: В ситуациях, где требуется оперативное и точное редактирование звука, таких как живые выступления или прямые трансляции, VST инструменты позволяют изменять звук непосредственно во время записи или воспроизведения, меняя высоту тона, частоту или другие эффекты.

2. Минимальная задержка: Для эффективной работы в реальном времени VST инструменты должны обрабатывать звук мгновенно. Это гарантирует отсутствие задержек или прерываний во время проигрывания, обеспечивая плавный и естественный опыт и сокращая время задержки.
3. Возможность взаимодействия с живыми инструментами и другим аудиооборудованием является ключевым преимуществом использования VST-инструментов. Артисты могут придать своим выступлениям глубину и гибкость, применяя VST-эффекты к инструментам в реальном времени, таким как гитара или клавишные.
4. Применение VST в прямых трансляциях и в играх не ограничивается только музыкальным производством. Пользователи могут использовать эти инструменты для потокового вещания, изменяя свои голоса в реальном времени во время общения или пения. Именно поэтому VST так популярны в стриминге и игровой индустрии, где звук можно мгновенно изменить, чтобы создать нужное настроение или соответствовать образам персонажей в играх.

Практический пример

Применение VST-плагинов для преобразования голоса с мужского на женский и наоборот позволяет оперативно изменять различные аспекты звучания, включая форманты и тональность. Это предоставляет возможность трансформировать голос, увеличивая верхние частоты и уменьшая высоту звука. Благодаря своей интерактивности, такие изменения могут быть внесены непосредственно во время исполнения или онлайн-трансляции. VST-плагины стали ключевым элементом арсенала современных творцов, будь то в звукозаписывающих студиях или на платформах для прямых эфиров и концертов, благодаря их способности эффективно функционировать в режиме реального времени. Эти инструменты открывают перед пользователями новые горизонты для манипуляции и настройки звука, что ранее было недостижимо.

Voice Modulator Plugins и Pitch Shifting Plugins

Обсуждая вопрос трансформации голоса от мужского к женскому или наоборот, нельзя не упомянуть популярные программные средства и приложения, которые используются как экспертами, так и начинающими. Инструменты для коррекции голоса и изменения его высоты активно применяются в разнообразных творческих сферах, включая музыкальное создание, онлайн-трансляции, разработку видеоигр и киноиндустрию, предоставляя гибкие и результативные методы для трансформации и улучшения звучания с помощью инновационных технологий [5–7]:

Плагины для коррекции голоса: Эти плагины служат ключевым инструментом для изменения голоса. Их основная функция заключается в корректировке и трансформации звучания голоса, что дает возможность полностью изменить его характер, делая его не узнаваемым. Эти приложения предоставляют обширный набор возможностей для преобразования мужского голоса в женский и наоборот, а также позволяют создавать необычные эффекты, такие как превращение голоса в звучание робота или монстра.

Проверенные средства для трансформации голоса:

- MorphVOX Pro: Этот инструмент пользуется огромной популярностью среди тех, кто стремится мгновенно изменить свою речь. Он предлагает обширный ассортимент голосовых эффектов, которые можно использовать для мужских и женских голосов. Особенностью MorphVOX Pro является возможность точной настройки голоса, чтобы он звучал натурально и подошел для разнообразных целей, включая игры и онлайн-трансляции [8].
- Clownfish Voice Changer: Этот инструмент отличается своей простотой и эффективностью. Он позволяет пользователям без труда менять свой голос, переходя от мужского к женскому или создавая необычные звуки, наподобие роботов или монстров [9].
- Voxal Voice Changer: Этот инструмент также обладает высокой степенью возможностей для изменения голоса. Он предлагает широкий спектр голосовых модификаций, которые можно применять в реальном времени. Особенностью Voxal Voice Changer является возможность точной настройки высоты тона и частоты [10].

Pitch Shifting Plugins: Сдвиг высоты тона (Pitch Shifting)

Этот метод позволяет корректировать высоту звучания голоса, при этом не влияя на его темп и продолжительность звучания, и является одним из наиболее распространенных способов обработки музыкальных и голосовых композиций. Применяя данную технику, можно поднять или понизить тональность голоса, чтобы он звучал более высоко (женственно) или низко (мужественно), при этом не меняя другие аспекты, вроде скорости или длительности звучания. Этот инструмент неocenим для трансформации голосов в соответствии с нужным полом или специфическими творческими запросами. Наиболее востребованными являются плагины для изменения высоты тона голоса [6–7]:

- Antares Auto-Tune: Один из наиболее известных инструментов для изменения высоты звучания голоса, применяемый как в музыке, так и для других коррекций звука. Широко используется в музыкальной сфере для исправления ошибок и из-

менения высоты звучания, но также представляет собой мощное средство для модификации голоса человека и его преобразования между полами.

- *Celemony Melodyne*: Профессиональный инструмент для точной настройки звука, который позволяет корректировать тон и изменять голос естественным и реалистичным образом. Обладает высокой степенью гибкости в управлении частотой звучания без утраты его первоначального качества [11].
- *Zynaptiq Pitchmap*: Позволяет быстро и точно регулировать высоту звучания и частоту, что делает его эффективным для плавного изменения звучания человеческого голоса с сохранением естественности [12].

Обсуждение

Применение инструментов

Во время прямых трансляций и записи подкастов, возможности модуляции голоса помогают изменять тембр и интонацию ведущих и дикторов, чтобы создавать различные характеры или атмосферу. Такие инструменты также применяются для трансформации голоса во время голосовых разговоров или общения в онлайн-чатах.

В игровой индустрии: Многие игроки используют подобные возможности, чтобы преобразить голос своего виртуального персонажа, придавая ему особенности, соответствующие его роли в игре, например, изменяя пол голоса с мужского на женский или наоборот.

В сфере музыкального творчества: Технологии изменения тона позволяют адаптировать голос исполнителя к мелодии, делая его более гармоничным или преобразуя его так, чтобы он звучал как голос другого пола, придавая новые оттенки творческому процессу.

Важность задержки и точности в редактировании звука

Задержка представляет собой один из критически важных параметров в области цифровой обработки звука, особенно при применении инструментов, таких как VST, для редактирования аудио в режиме реального времени. Понятие задержки определяется как интервал времени между моментом поступления звукового сигнала (например, речи или музыки) и временем, когда измененный звук воспринимается слушателем. Увеличение задержки приводит к ухудшению плавности и реализма процесса обработки звука, что негативно сказывается на эффективности и качестве окончательного результата [13].

Фонетический анализ:

1. Основные частоты: Основные частоты — это низкочастотные составляющие, определяющие высоту человеческого голоса, и они являются основным отличием между мужскими и женскими голосами. У мужчин основные частоты обычно ниже (от 85 до 180 Гц), тогда как у женщин они выше (от 165 до 255 Гц). Эта разница обусловлена физиологическими различиями в голосовых связках и размерах голосового аппарата. При использовании VST-плагинов для изменения голоса с мужского на женский или наоборот, основные частоты корректируются — понижаются при преобразовании женского голоса в мужской и повышаются при преобразовании мужского голоса в женский. Это достигается с помощью инструментов, таких как Pitch Shifting, которые контролируют диапазон частот голоса для создания более соответствующего типа звука.
2. Частотные диапазоны (форманты) и их изменение: Форманты — это набор звуковых частот, определяемых формой и размером ротовой полости и гортани. Форманты придают звуку его характерную «окраску», позволяя различать звуки одинаковой высоты тона. В мужском голосе форманты ниже, чем в женском, из-за большего размера голосовой полости у мужчин. При изменении голоса между полами с использованием VST-плагинов форманты точно изменяются. Например, при преобразовании мужского голоса в женский требуется повышение формант, чтобы голос звучал более остро и женственно. Некоторые продвинутые инструменты, такие как Melodyne или Antares Auto-Tune, позволяют модифицировать форманты независимо от тона, что дает возможность изменять частотные диапазоны, не влияя на исходную частоту звука.
3. Высота тона (Pitch): Высота тона — это основной параметр, определяющий, как высоко или низко звучит голос. Женский голос обычно имеет более высокую высоту тона, чем мужской. Корректировка высоты тона — один из ключевых аспектов изменения голоса между полами. С помощью таких плагинов, как Pitch Shifter или Auto-Tune, тон можно повысить или понизить в зависимости от требуемого результата.
 - Для преобразования мужского голоса в женский тон повышается на несколько октав, чтобы голос звучал выше и ближе к женскому голосу.
 - Для преобразования женского голоса в мужской тон понижается, чтобы сделать голос глубже и тяжелее, что характерно для мужского голоса. Процесс изменения тона требует точности, чтобы сохранять реалистичность голоса и избегать искусственности и искажений. Продвинутые инструменты помогают достичь баланса между изменением тона и сохранением качества звука.

4. Гибкость голоса: Гибкость голоса — это способность легко изменять высоту и тональность во время разговора или пения. Гибкость голоса заметно отличается между полами: женский голос обычно более гибок в манипулировании высокими нотами, тогда как мужской голос более стабилен и имеет меньший диапазон частотных изменений.
- Женский голос: обладает большей гибкостью в манипуляции высокими частотами, легко переходя от низких к высоким тонам, что придает ему живость и выразительность.
 - Мужской голос: обычно менее гибок, поддерживая относительную стабильность в высоте тона и частотах, с заметно меньшей способностью переходить между нотами.

Для анализа можно использовать преобразование Фурье, искусственный интеллект с машинным обучением или глубокое обучение также используются в области голоса, но некоторые методы используются по-другому, некоторые создают новые голоса, некоторые полностью изменяют голос, а другие добавляют эффекты в зависимости от анализа искусственного интеллекта и запросов пользователей

Заключение

Технологии редактирования звука с использованием VST и других программ стали ключевыми в цифровом творчестве, таких как кинопроизводство и игры. Эти инструменты позволяют художникам и создателям изменять голоса персонажей так, чтобы они соответствовали сюжетным задачам, усиливая погружение в контент. В играх они помогают игрокам изменять голоса своих персонажей, повышая уровень взаимодействия с виртуальным миром. Несмотря на значительный прогресс в технологиях редактирования звука, существуют области для дальнейших исследований. Одной из главных задач является улучшение качества изменённых голосов для повышения их реалистичности, а также минимизация искажений. Ещё одной важной областью является снижение задержки, чтобы обеспечить мгновенное редактирование звука без задержек. Также активно развиваются технологии искусственного интеллекта, которые позволяют этим инструментам более точно понимать контекст звука и естественно изменять голос.

ЛИТЕРАТУРА

1. Танев Г., Божиновский А. Технология виртуальной студии и ее применение в производстве цифровой музыки. — 2013.
2. Танев Г., Божиновский А. Технология виртуальной студии в музыкальном производстве // Международная конференция по инновациям в области ИКТ. — Гейдельберг: Springer International Publishing, 2013. — С. 231–241.
3. Осорио-Гонага Р. Проектирование и реализация цифровых фильтров в архитектуре виртуальной студии Steinberg (vst) // Съезд 119 Общества аудиоинженеров. — Общество аудиоинженеров, 2005.
4. Ким Д.Х. и др. Снижение шума при распознавании речи с использованием технологии виртуальной студии // Материалы конференции Корейского общества обработки информации. — Корейское общество обработки информации, 2018. — С. 21–24.
5. Писански К. и др. Модуляция голоса: окно в истоки голосового контроля человека? // Тенденции когнитивных наук. — 2016. — Т. 20. — №. 4. — С. 304–318.
6. Рахман С. Изменение высоты голоса в реальном времени // Компьютерная техника. — Т. 2. — С. 35163.
7. Чжэн Ю. и др. Вокальная привлекательность и произвольное изменение высоты голоса // Эволюция и поведение человека. — 2020. — Т. 41. — №. 2. — С. 170–175.
8. Ли Х. и др. Звук галлюцинаций: к более убедительной имитации внутренних голосов // Материалы конференции CHI 2022 года по человеческому фактору в вычислительных системах. — 2022. — С. 1–11.
9. Колли О. и др. Межвидовая изменчивость криков рыб-клоунов: степень сходства у близкородственных видов // Эволюционная биология БМК. — 2011. — Т. 11. — С. 1–9.
10. Анеджа П. Статистический анализ и тестирование надежности различных акустических элементов при судебно-медицинской экспертизе преобразованных образцов голоса // Доступно по адресу SSRN 4466886.
11. Бриджес-младший СР. Влияние программ настройки программного обеспечения на записи вокала // Съезд 149 Общества аудиоинженеров. — Общество аудиоинженеров, 2020.
12. Ассортимент Н.М.Б. Интересующие продукты // Компьютерный музыкальный журнал. — 2016. — Т. 40. — №. 1. — С. 76–87.
13. Бородин А.М., Александрова Н.В. ИЗМЕРИТЕЛЬ СКОРОСТИ ЗВУКА В ЖИДКОЙ СРЕДЕ. — 2005.
14. Мицкевич М.В. СОЗДАНИЕ МУЗЫКИ С ИСПОЛЬЗОВАНИЕМ ВИРТУАЛЬНЫХ VST-ИНСТРУМЕНТОВ // Региональная информатика (РИ-2022). — 2022. — С. 327–329.
15. Аракава Р. и др. Цифровая речевая коррекция: измененная слуховая обратная связь на основе преобразования голоса для трансформации саморепрезентации // Материалы Международной конференции по мультимодальному взаимодействию 2021 года. — 2021. — С. 159–167.