АУДИОПЛЕЕР С голосовым управлением для незрячих

Гасанов Э.В. (elearn@yandex.ru), Абрамова Н.А. (ninabramov@yandex.ru),

Миронов В.В. (vovamironov2900@gmail.com)

Московский политехнический университет, г. Москва,

АНОО «Школа Сосны», Московская область, Одинцовский район

Аннотация

В работе рассмотрены возможности российского одноплатного компьютера Repka Pi 3 для создания аудиоплеера с голосовым управлением. Аудиоплеер с голосовым управлением, благодаря использованию голосовых команд, существенно облегчает незрячим людям использование аудиоплеера для прослушивания аудиокниг. Проект выполнен в рамках программы подготовки инженеров будущего.

По оценкам исследователей, в 2020 году лишены зрения 43,3 млн людей по всему миру. Еще 295 млн имеют умеренные или тяжелые нарушения зрения. Исследователи прогнозируют, что к 2050 году численность слепых людей достигнет 61 млн, с нарушениями зрения – 474 млн. [1].

Ученые связывают увеличение числа незрячих и слабовидящих людей с возрастающими нагрузками на зрение – прежде всего за счет чрезмерного использования компьютеров и мобильных телефонов.

Инвалидам по зрению, визуальное восприятие которых ограничено болезнью или полной слепотой, для развития и получения образования, а также востребованных рынком труда профессий и навыков необходимы альтернативные возможности и источники получения информации.

Внедрение разнообразных ассистивных технологий — одна из таких возможностей, которые позволяют улучшить качество жизни людей с ограниченными возможностями.

К ассистивным технологиям относятся устройства, а также программные и иные средства, использование которых позволяет расширить возможности людей с ограниченными возможностями для доступа к информации, процессу общения, информационно-образовательной среде.

Аудиоплеер с голосовым предоставляет незрячим людям возможность управлять аудиоплеером для прослушивания аудиокниг с помощью голосовых команд.

В ГОСТ Р 59224— 2020 «ЦИФРОВАЯ «ГОВОРЯЩАЯ» КНИГА ДЛЯ СЛЕПЫХ И СЛАБОВИДЯЩИХ» аудиокнига для незрячих обозначается термином цифровая «говорящая» книга: Электронная аудиокнига, записанная в цифровом криптозащищенном аудиоформате, предназначенная для прослушивания с помощью специальных программных и аппаратных тифлотехнических средств. Специальное устройство для чтения «говорящих» книг на флеш-картах с озвученным интерфейсом пользователя, предназначенное для людей с нарушенной функцией зрения (слепых или слабовидящих) называется тифлофлешплеером [2].

На сегодняшний день для незрячих людей существует множество разнообразных тифлофлешплееров для прослушивания цифровых «говорящих» книга. Управление осуществляется с помощью выполнения заданной последовательности нажатий на определенные кнопки. Ни один из представленных тифлофлешплееров не имеет функции управления с помощью голосовых команд. В данной работе предложена отличная от других решений реализация аудиоплеера для незрячих: использование голосовых команд для управления аудиоплеером. Такой способ управления существенно облегчит незрячим использование аудиоплеера при прослушивании аудиокниг.

Также прослушивание незрячими аудиокниг возможно с помощью виртуальных голосовых ассистентов, таких как Apple Siri, Google Assistant, Amazon Alexa, Microsoft Cortana, Алиса, Маруся, Салют и другие.

Однако у всех подобных систем есть два существенных минуса. Первый — это необходимость в постоянном интернет-подключении. Второй минус — это цена. Услуги виртуальных голосовых ассистентов являются платными. А для доступа к аудиокнигам необходима дополнительная платная подписка.

Разработанный в рамках данного проекта прототип аудиоплеера с голосовым управлением для прослушивания аудиокниг в формате MP3 незрячими людьми работает без доступа к Интернет и позволяет использовать аудиокниги в формате MP3 из свободных бесплатных онлайн-библиотек.

Основной задачей при работе над проектом по созданию аудиоплеера была разработка электронного устройства для прослушивания аудиокниг и программного обеспечения для управления этим устройством.

Одной из задач, которую нужно было решить при разработке программного обеспечения - это распознавание речи.

Распознавание речи (Speech recognition) — это огромная область компьютерных наук и компьютерной лингвистики, которая разрабатывает методы и технологии, позволяющие распознавать разговорную речь и переводить её в цифровую информацию, например, в текст.

Подобный процесс также часто называют «автоматическим распознаванием речи», или сокращённо ASR (Automatic Speech Recognition). Технологию преобразования речи в текст часто называют просто «Speech-to-Text», или сокращённо STT [3].

В настоящее время разработано множество систем распознавания речи, которые предоставляют библиотеки (API, application programming interface) для разработки приложений голосового управления. Для разработки программного обеспечения устройства прослушивания аудиокниг библиотека должна соответствовать следующим критериям:

1) поддерживать русский язык;

2) работать без доступа к Интернет;

3) сделана для потоковой обработки звука, что позволяет реализовать мгновенную реакцию на команды.

Этим критериям удовлетворяет библиотека Vosk, разработанная российской компанией АЦ Технологии. Библиотека Vosk работает на устройствах Android, iOS, Raspberry Pi, серверах и поддерживает несколько популярных языков программирования – Python, Java, C#, Javascript.

Поэтому в качестве контроллера управления (вычислительного модуля) устройством прослушивания аудиокниг был выбран российский одноплатный компьютер Repka Pi 3, который является полноценной альтернативой западному аналогу Raspberry Pi 3 (самый популярный одноплатный компьютер). Repka Pi 3 работает под управлением собственной операционной системы Repka Pi, реализованной на основе портированной операционной системы Ubuntu (дистрибутив Linux) [4].

Для ввода голосовых команд и прослушивания аудиофайлов использована компьютерная гарнитура (наушники и микрофон), которая подключена к Repka Pi 3 с помощью внешней звуковой карты Creative Sound Blaster Play! 3.

Для голосового управления устройством прослушивания аудиокниг было разработано программное обеспечение на языке Python с использованием библиотеки системы распознавания речи Vosk.

В настоящее время голосовое управление устройством реализовано с помощью следующих голосовых команд: «старт», «стоп», «вперед», «назад», «закладка», «начни с закладки», «выключить».

Аудиоплеер с голосовым управлением, благодаря использованию голосовых команд, существенно облегчает незрячим людям использование аудиоплеера для прослушивания аудиокниг.

Литература

1. 1. Представлен глобальный анализ распространенности слепоты и нарушений зрения [Электронный ресурс] URL: <https://medvestnik.ru/content/news/Predstavlen-globalnyi-analiz-rasprostranennosti-slepoty-i-narushenii-zreniya.html> (дата обращения 19.02.2024)
2. 2. ГОСТ Р 59224— 2020 «ЦИФРОВАЯ «ГОВОРЯЩАЯ» КНИГА ДЛЯ СЛЕПЫХ И СЛАБОВИДЯЩИХ» [Электронный ресурс] URL: <https://docs.cntd.ru/document/1200177434> (дата обращения 19.02.2024)
3. 3. Как сделать робота на ROS своими руками. Часть 3: распознавание речи для голосового управления [Электронный ресурс] URL: <https://amperka.ru/blogs/projects/abot-robot-part-3> (дата обращения 19.02.2024)
4. 4. Российский одноплатный компьютер Repka Pi 3 [Электронный ресурс] URL: <https://repka-pi.ru> (дата обращения 19.02.2024)