

ИСПОЛЬЗОВАНИЕ ВОЗМОЖНОСТЕЙ РЕЧЕВОЙ ПЛАТФОРМЫ MICROSOFT SPEECH PLATFORM ПРИ РАЗРАБОТКЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ЛЮДЕЙ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ

Введение. Развитие информационных технологий в настоящее время повлияло на все сферы деятельности человека. Сейчас невозможно представить человека, у которого бы не было ноутбука или персонального компьютера. Зачастую все операции на компьютере требуют ввода информации путем манипулирования устройствами ввода, такими как клавиатура, мышь и др. Однако существует такая категория граждан, как лица с ограниченными возможностями здоровья. Зачастую в программных продуктах никак не предусмотрен ввод и управление программой голосом.

В данной исследовательской работе будут рассмотрены возможности речевой платформы Microsoft Speech Platform.

Основная часть. В данной исследовательской работе основной целью является показ того, что при разработке программного обеспечения на языке C# с использованием программной платформы .NET Framework можно использовать голосовое управление и ввод, что, в свою очередь, упростит пользование программными продуктами людям с ограниченными возможностями здоровья — незрячим (слабовидящим).

Основными задачами исследовательской работы являются: описать возможности речевой платформы Microsoft Speech Platform; разработать простейшее приложение с голосовым интерфейсом; изучить возможности платформы Microsoft Speech Platform SDK; написать приложение с использованием языка программирования C# и программной платформы .NET Framework.

Microsoft Speech Platform — это набор программных модулей и средств разработки, позволяющих разработчикам создавать приложения и сервисы с поддержкой речевых технологий (распознавание речи и синтез речи по тексту), а пользователям взаимодействовать с такими приложениями. В официальных анонсах можно встретить указание на то, что речевая платформа является в определённой степени развитием речевых технологий Microsoft, известных под общим названием Microsoft Speech API (MS SAPI), однако это вполне самостоятельный программный продукт, который можно установить в системе независимо от того, присутствует ли в ней MS SAPI или нет [1].

Речевая платформа включает в себя компоненты времени выполнения Microsoft Speech Platform Runtime, которые необходимо установить всем желающим использовать платформу; средства разработки речевых приложений и служб Microsoft Speech Platform SDK, которые необходимы исключительно разработчикам программного обеспечения [2]; а также языковые модули Language Runtime, устанавливаемые по выбору пользователя. Сейчас в речевой платформе версии 11.0 поддерживаются 26 языков (в их числе русский и чешский).

Для знакомства с Microsoft Speech Platform мы напишем простое приложение, в котором произвольно будут распознаваться запрограммированные слова.

Речевая платформа Microsoft Speech Platform предназначена для распознавания и синтеза речи.

Распознавание речи — процесс преобразования произнесённых слов в печатный текст. Распознавание речи включает в себя: захват и оцифровку звука, произнесённого в микрофон; преобразование оцифрованного звука в фонемы; конструирование из фонем слов; анализ контекста, в котором было произнесено слово, и (при необходимости) замена слов на похожие по звучанию.

Распознающая программа (движок распознавания текста) итерационно сравнивает распознанный текст с правилами грамматики приложения и при совпадении текста с серией правил формирует выходной поток XML с использованием языка разметки семантики. Выходной поток содержит распознанный текст, значения вероятностей правильного распознавания и может содержать семантические значения, присвоенные с помощью разметки интерпретации семантики. Распознанный текст обычно используется для ввода данных под диктовку и для управления приложениями с помощью речевых команд.

Синтез речи — процесс преобразования текста в произносимые слова. Синтез речи включает в себя: разделение слов на фонемы; нахождение текста, требующего преобразование в символы, такого как числа, количество валюты и пунктуация; генерация цифрового звука для воспроизведения.

Движки преобразования текста в речь могут использовать один из двух способов синтеза голоса: генерировать звуки, подобные издаваемым голосовыми связками, и использовать различные фильтры для моделирования длины горла, формы ротовой полости, расположения губ и языка; собирать его из многочисленных кусков речи, записанных диктором.

В данной работе было разработано простейшее приложение для детектирования голосовых команд. На примере Windows Form приложения было обработано событие Show главной формы. Таким образом, когда приложение открывается, начинает работать детектор. На рисунке 1 представлено главное окно приложения, содержащее на форме один элемент пользовательского интерфейса — Label.



Рисунок 1 — Главное окно приложения

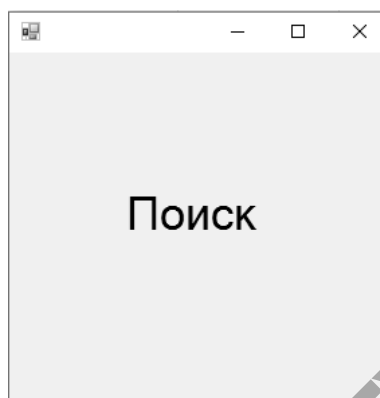


Рисунок 2 — Пример распознавания

Приложение запрограммировано таким образом, что выводит распознанные команды в Label. Если начать произносить текст, а детектор встретит запрограммированные слова или фразы, он запишет их в Label. Пример распознанной команды и ее вывода в Label представлен на рисунке 2.

Это достаточно простой пример того, как настроить голосовое управление в программе. Таким же образом можно реализовать, например, нажатие на кнопку, закрытие программы, т. е. обработать практически любое событие.

Заключение. В работе предложен метод управления программами с использованием речевой платформы Microsoft Speech Platform. Эксперименты показали, что предложенный метод позволяет с достаточно высокой точностью распознавать голосовые команды, произнесенные человеком, а также с помощью их обрабатывать события элементов пользовательского интерфейса Windows Form приложений.

На примере простейшего приложения показано, как можно управлять программами с помощью голоса. Такая полезная функция может стать незаменимым помощником для незрячих (слабовидящих) людей, работающих с программными продуктами. Кроме распознавания речи данная платформа может также синтезировать речь, т. е. в программе можно реализовать аудиоподсказки.

Список цитируемых источников

1. Речевая платформа Microsoft — Runtime (версия 11) [Электронный ресурс]. — 2018. — Режим доступа: <https://www.microsoft.com/en-us/download/details.aspx?id=27226>. — Дата доступа: 06.09.2018.
2. Microsoft Speech Platform — комплект для разработки программного обеспечения [Электронный ресурс]. — 2018. — Режим доступа: <https://www.microsoft.com/en-us/download/details.aspx?id=27225>. — Дата доступа: 06.09.2018.
3. *Томашенко, Н. А.* Исследование проблемы сбалансированности данных при построении акустических моделей систем автоматического распознавания речи [Электронный ресурс]. — 2018. — Режим доступа: <https://cyberleninka.ru/article/n/issledovanie-problemy-sbalansirovannosti-dannyh-pri-postroenii-akusticheskikh-modeley-sistem-avtomaticheskogo-raspoznaniya-rechi>. — Дата доступа: 06.09.2018.

УДК 004.932

Е. Г. Шапович¹, В. В. Кравчяня²

¹Учреждение образования «Барановичский государственный университет», Барановичи,

²ЧПТУП «Сансет», Барановичи

СОВРЕМЕННЫЕ ВИРТУАЛЬНЫЕ ЭКСКУРСИИ И СРЕДСТВА РАЗРАБОТКИ ВИРТУАЛЬНЫХ ЭКСКУРСИЙ

Введение. В настоящее время различные объекты, такие как музеи, отели, университеты и т. д., используют виртуальную реальность для создания виртуальных туров. Виртуальные туры имитируют нахождение человека на этих объектах, дают возможность просматривать все экспозиции в любое удобное время и решать проблемы доступности, так как не каждый из нас сможет посетить особенно удаленные объекты по тем или иным причинам.

Виртуальная экскурсия (тур) считается способом реалистичного отображения трехмерного многоэлементного пространства на экране. Элементы подобных проектов представляют собой сферические панорамы, которые связаны между собой посредством интерактивных переходов, например хотспота. Простыми словами,