

ГОЛОСОВОЕ УПРАВЛЕНИЕ ПЕРСОНАЛЬНЫМ КОМПЬЮТЕРОМ

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Кива В.С.

Лукьянец В.Г. – к.т.н., доцент

Статья посвящена характеристике сути голосового управления и возможности применения этой технологии на персональном компьютере. Рассмотрены основные реализованные алгоритмы распознавания речи.

В настоящее время проблема разработки и создания эффективных систем распознавания речи и голосового управления персональным компьютером имеет большую актуальность. Совместное использование таких систем является основой полнофункционального голосового интерфейса, объем применения которого, очень широк на практике. Исследования в области интерфейса речи ведутся многими учеными, а разработка осуществляется крупнейшими компьютерными организациями, включая Microsoft, Intel и IBM. Использование распознавания команд, посредством голосового управления на персональном компьютере, может иметь широкий диапазон и может быть реализовано во многих областях общества и промышленности.

На современном этапе развития технологий трудно представить полный переход к голосовому управлению, что станет возможно только тогда, когда персональный компьютер научится полностью понимать человеческую речь, а не просто ограниченный набор команд. Но наличие режимов голосового управления в различных приложениях и операционных системах действительно актуально, потому что не все люди имеют возможность свободно использовать свои руки. В принципе, управление голосом с помощью компьютера упрощает работу с ним, потому что команды передаются быстро и без усилий через микрофон.

Ведущие компании по разработке программного обеспечения, разработке операционных систем и внедрению новых высокотехнологичных устройств понимают все преимущества включения голосового управления в список возможностей своих продуктов, поэтому в этом направлении ведется интенсивная работа. Среди операционных систем, созданных Microsoft, функция голосового управления доступна только в версии Windows 8. Те же функции поддерживаются операционными системами, созданными Apple, особенно благодаря применению Siri [1].

Чтобы настроить голосовые команды в операционной системе, на персональном компьютере должна быть установлена соответствующая программа. Взаимодействие с компьютером может быть создано не только с помощью клавиатуры и мыши. Теперь стали доступны и голосовые команды. Существует специальный инструментарий, который позволяет вам это делать. Его функция включает в себя не только написание текста для диктовки или декодирование аудиозаписей. Через голосовые команды можно запускать приложения и использовать их [2].

Основой для создания систем голосового управления являются алгоритмы и методы распознавания речи.

Наиболее распространенными и эффективными являются следующие методы:

1. MFCC (Mel-Frequency Cepstral Coefficients) – это метод кепстральных коэффициентов, который заключается в вычислении коэффициентов спектра Фурье, наложения на полученный спектр набора фильтров шкалы мел, выполнения логарифмирования измененного спектра и реализации дискретного косинусного преобразования.
2. LPCC (Linear Predictive Cepstral Coefficients) – это метод кепстральных коэффициентов линейного предсказания, который основывается на вычислении коэффициентов авторегрессионной модели для каждого фрейма аудио сигнала.
3. PLP (Perceptual Linear Predictive) – это метод коэффициентов перцептивного линейного предсказания, который отличается от метода LPCC тем, что учитываются особенности восприятия различных частот человеком - перед вычислением параметров авторегрессионной модели сигнал проходит определенную предобработку. Вычисленный мгновенный спектр Фурье преобразуется в спектр на шкале барков, после чего выполняется операция свертки маскирующих кривых критических полос с полученным спектром для получения эффекта маскировки частоты. Далее производится аппроксимация кривой громкости и кепстральная обработка [3].

Признанные методы анализа приведенных методов показали, что большинство усилий разработчиков сосредоточено на частотных характеристиках выделения человеческого вокального канала и, следовательно, на качественных методах, основанных на обработке кепстрального сигнала.

Настройка системы распознавания речи подразумевает способность управлять шумодавлением, определять время задержки и результаты распознавания, а также, процент слияния нескольких результатов распознавания. Итоговый результат возникает в окне и позволяет редактировать речевую базу без специального знания системы. Настройки распознавания голоса на персональном компьютере позволяют настроить следующие параметры: автозапуск с системой Windows; ввод журнала событий и ошибок и сохранение его в файле; определить, как сохранить журнал ошибок и событий и управлять виджетами [4].

Программное обеспечение и виджеты для голосового управления реализуются в Microsoft Visual Studio с поддержкой интегрированной среды разработки Windows Forms и WPF с использованием Visual C #, являются реализацией языка C# и предназначены для демонстрации основных функций речевого модуля Microsoft SpeechAPI.

Анализ рынка программ показывает, что лидером среди программ, которые позволяют управлять компьютером на русском языке, по качеству и точности распознавания является программа Tuple. Есть, также такие программы как: RealSpeaker, Web Speech, Горыныч, и т.д., но все они менее точны или у них не полная интеграция с ОС.

Таким образом, в результате исследования были рассмотрены и проанализированы наиболее успешные методы распознавания речевых команд и имеющееся эффективное программное обеспечение для управления голосом на персональном компьютере.

Список использованных источников:

1. Самые совершенные программы управления голосом [Электронный ресурс]. – Режим доступа: https://www.gkh11.ru/news/samye_overshennye_programmy_upravlenija_golosom/2016-01-28-1770. – Дата доступа: 24.03.2018.
2. Управление голосом и жестами на компьютере [Электронный ресурс]. – Режим доступа: <https://bursin.ru/upravlenie-golosom-i-zhestami-na-kompyutere/>. – Дата доступа: 22.03.2018.
3. Кравченко, К.В. Автоматизированная система голосового русскоязычного управления операционной системой Windows / К.В. Кравченко, Р.А. Дьяченко // Современные проблемы науки и образования. – 2014. – № 3.
4. Как управлять компьютером при помощи голосовых команд? [Электронный ресурс]. – Режим доступа: <http://nastroyse.ru/programs/review/upravlenie-kompyuterom-s-pomoshhyu-golosa.html>. – Дата доступа: 23.03.2018.

ПРОГНОЗИРОВАНИЕ КЛИМАТИЧЕСКИХ УСЛОВИЙ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Клещиков А.С.

Медведев С.А. – к.т.н., доцент

Влияние прогноза климатических условий для общества, бизнеса, сельского хозяйства заставляет уделять большое внимание данному вопросу. Последние несколько лет показывают потрясающие достижения в области прогнозирования климата. Все системы и методы, разработанные до сих пор, используют температуру поверхности моря в качестве основного фактора, используемого для прогноза. Затем уже используются статистические и математические модели для дальнейшего прогноза климата. Однако есть и другой подход к прогнозированию климатических условий, который основывается на исторических метеорологических показателях региона, таких как осадки, скорость ветра, точка росы, температура и т.д.

Основная цель данной статьи - как использовать технологии интеллектуального анализа данных (в частности, алгоритма К-ближайших соседей), как разработать систему, которая использует исторические данные для прогнозирования климатических условий заданного региона, города или страны на месяцы вперед.

Интеллектуальный анализ данных - это недавняя разработка в области очень больших баз данных и хранилищ данных. Он используется для обнаружения скрытых полезных паттернов в огромных базах данных. Интеллектуальный анализ данных может быть классифицирован по его методам на три основных типа: поиск ассоциативных правил, кластерный анализ и классификация/предсказывание.

Алгоритм К-ближайших соседей – это алгоритм классификации, который основывается на формуле дистанции в евклидовом пространстве. Данная формула используется для выяснения близости между неизвестными образцами с известными классами. Неизвестный образец присваивается тому классу, который является наиболее распространенным среди К соседей данного образца, классы которых уже известны.

Функция вычисления дистанции в евклидовом пространстве выглядит следующим образом:

$$d(p, q) = \sqrt{\sum_{k=1}^n (p_k - q_k)^2} \quad (1)$$

Шаги алгоритма К-ближайших соседей выглядят следующим образом:

вычисление расстояния между новым образцом и известными образцами по формуле (1);

сортировка полученных расстояний так, чтобы $d_i \leq d_{i+1}$ и выборка К образцов с наименьшими значениями расстояний;

новому образцу присваивается класс с наибольшим количеством элементов в выборке.

Для применения алгоритма К-ближайших соседей для прогнозирования климатических условий необходимо указать четыре параметра: текущая дата (дата, от которой будет отсчитываться прогноз) (CD), количество дней для прогноза (N), значение К алгоритма и атрибуты для прогноза.

Затем, для вычисления расстояния в евклидовом пространстве, исходный массив данных разбивается на последовательности. Каждая последовательность имеет размер S, который рассчитывается как произведение количества дней прогноза N на количество атрибутов прогноза. Базовой последовательностью являются записи за предыдущие N дней для выбранных атрибутов. Общее количество последовательностей (T) вычисляется как общее количество записей делить на количество записей в базовой последовательности. Дальнейшее вычисление расстояний можно выразить с помощью псевдокода:

```
While i < Общее_Количество_Записей  
For j = 0 to S
```