

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СОВРЕМЕННЫХ СИСТЕМ РАСПОЗНАВАНИЯ РЕЧИ

Шубный А.В.,
научный руководитель доц. Медведев М.С.
Сибирский федеральный университет

В данном докладе представлен ряд наиболее доступных современных систем распознавания речи, а также их сравнительные характеристики.

VoiceNavigator (компания «Центр речевых технологий» (ЦРТ))

Это высокотехнологичное решение для контакт-центров, предназначенное для построения Систем Голосового Самообслуживания (СГС). VoiceNavigator позволяет автоматически обрабатывать вызовы с помощью технологий синтеза и распознавания речи.

Характеристики:

- дикторонезависимость;
- устойчивость к окружающим шумам и помехам в телефонном канале;
- распознавание русской речи работает с надежностью 97% при на словаре в 100 слов;

VoiceNavigator Web

Навигация по веб-ресурсам при помощи голоса. Разработка компании «Центр речевых технологий», позволяет управлять навигацией сайта при помощи речевых команд. Характеристики данной системы распознавания идентичны характеристикам VoiceNavigator.

Speereo Speech Recognition (SSR) (Российская ИТ компания «Speereo Software»)

Разработка приложений в индустрии мобильных устройств, а также для РС. Использование приложений в веб-пространстве. Распознавание речи происходит непосредственно на устройстве, а не на сервере, что является ключевым преимуществом, по мнению разработчиков.

Характеристики:

- распознавание русской речи работает с надежностью около 95%;
- дикторонезависимость;
- словарный запас около 150 тыс. слов;
- одновременная поддержка нескольких языков;
- компактный размер движка.

Sakrament ASR Engine (разработка компании «Сакрамент»)

Осуществляет высокоточное распознавание речи на различных платформах. Технология распознавания речи используется при создании средств речевого управления – программ, управляющих действиями компьютера или другого электронного устройства с помощью голосовых команд, а также при организации телефонных справочных и информационных служб.

Характеристики:

- дикторнезависимость;
- языконезависимость;
- точность распознавания достигает 95-98%;
- распознавание речи в виде выражений и небольших предложений;
- нет возможности обучения.

Google Voice Search (компания «Google»)

Голосовой поиск от компании Google. Ранее поиск применялся исключительно в мобильных устройствах. С недавнего времени голосовой поиск от Google встроен в браузер Google Chrome, что позволяет использовать этот сервис на различных платформах.

Характеристики:

- поддержка русского языка;
- возможность встраивать распознавание речи на веб-ресурсы;
- голосовые команды, словосочетания;
- для работы необходимо постоянное подключение к сети internet.

Dragon NaturallySpeaking 11 (компания «Nuance»)

Мировой лидер в программном обеспечении по распознаванию человеческой речи. Возможность создавать новые документы, отправлять электронную почту, управлять популярными браузерами и разнообразными приложениями посредством голосовых команд.

Характеристики:

- отсутствует поддержка русского языка;
- точность распознавания до 99%.

ViaVoice (компания «IBM»)

ViaVoice представляет собой программное ядро для аппаратных реализаций. Компания ProVox Technologies на основе этого ядра создала систему для диктовки отчетов врачей-радиологов **VoxReports** .

Характеристики:

- точность распознавания достигает 95-98%;
- дикторнезависимость;
- словарь системы ограничен набором специфических терминов.

Рассмотрев сравнительные характеристики современных систем распознавания речи, можно сделать следующие выводы:

- наиболее перспективными разработками на российском рынке являются продукты компании «Центр речевых технологий» (ЦРТ);
- большинство популярных на сегодняшний день систем распознавания работают с изолированными словами;
- разработчики достигли высокой точности в командных системах (речевые интерфейсы, голосовое управление);
- в наиболее распространенных современных приложениях точность распознавания составляет в среднем 95-99% ;
- задача распознавания слитной речи в достаточной степени не решена.

Результаты сравнительного анализа систем распознавания речи представлены в таблице 1.

Таблица 1 – Сводная таблица характеристик систем распознавания речи

| Название системы | Назначение | Структурная единица | Обучение | Поддержка русского языка | Дикторо независимость | Коэфф. расп. |
|------------------------------------|--|---------------------|----------|--------------------------|-----------------------|--------------|
| VoiceNavigator, VoiceNavigator Web | Командная система, речевой интерфейс, распознавание речи | Слово, фраза | Да | Да | Да | 97% |
| Speereo Speech Recognition | Голосовое управление | Слово | Нет | Да | Да | 95% |
| Sakrament ASR Engine (Сакрамент) | Речевой интерфейс | - | Нет | Да | Да | 95% |
| Google Voice Search | Голосовой поиск | Слово, фраза | Нет | Да | Да | - |
| Dragon NaturallySpeaking | Голосовое управление, распознавание речи | - | Нет | Нет | Нет | 98% |
| ViaVoice, VoxReports | Голосовое управление, распознавание речи | - | Нет | Нет | Да | 95% |