

```
# Версию Ruby следует указывать в конце файла  
ruby "2.1.2"
```

Обратите внимание, вы должны использовать эту же версию Ruby и у себя.

Если вы используете web-сервер отличный от WEBrick (Н-р: Thin), что, кстати, рекомендуется самим сервисом Heroku, то вам необходимо будет создать так же Procfile приблизительно следующего содержания:

```
web: bundle exec thin start -p $PORT
```

ЛИТЕРАТУРА

Heroku [Электронный ресурс] / Wikipedia. – 2017. / Режим доступа:

<https://ru.wikipedia.org/wiki/Heroku/>. – Дата доступа: 25.03.2017

Развертываем свой сайт на Heroku [Электронный ресурс] / Habrahabr. – 2017. / Режим доступа: <https://habrahabr.ru/post/232679/>. – Дата доступа: 25.03.2017

Cloud Application Platform [Электронный ресурс] / Heroku. – 2017. / Режим доступа: <https://www.heroku.com/> – Дата доступа: 25.03.2017

УДК 004.5(075.8)

Студ. А. А. Етко

Науч. рук. ст. преп. Т. В. Кишкурно
(кафедра информатики и веб-дизайна, БГТУ)

СЕРВИСЫ РАСПОЗНАВАНИЯ И СИНТЕЗА РЕЧИ ПЛАТФОРМЫ IOS

Актуальность набора текста голосом на сегодняшний день очевидна. При разработке приложения для смартфона в виде чата, в который входит: поиск контактов, добавление контактов, создание одиночных и групповых чатов, обмен сообщениями – было принято решение использовать возможность управления программным средством и ввода сообщения – голосом.

Реализация своей собственной системы распознавания речи является очень сложной, трудоемкой и ресурсозатратной задачей, которую сложно выполнить в рамках данной работы. Поэтому предполагается интегрировать представленную технологию видео-

идентификации в системы распознавания речи, которые имеют для этого специальные возможности.

В качестве платформы для создания приложения выбрана операционная система iOS, которая разработана американской компанией Apple – для смартфонов, электронных планшетов и носимых проигрывателей, первоначально – для iPhone и iPod touch, позже – для таких устройств, как iPad и Apple TV.

Эта платформа предоставляет качественные сервисы распознавания и синтеза речи и отличается от других платформ, наличием голосового помощника Siri, который стремительно развивается, постоянно увеличивая базу языков. Распознавание речи – процесс преобразования речевого сигнала в цифровую информацию (например, текстовые данные). Обратной задачей является синтез речи.

Siri – облачный персональный помощник и вопросно-ответная система, программный клиент которой входит в состав iOS компании Apple. Данное приложение использует обработку естественной речи, чтобы отвечать на вопросы и давать рекомендации. Siri приспосабливается к каждому пользователю индивидуально, изучая его предпочтения в течение долгого времени.

Sirikit поддерживает ряд сценариев, каждый из которых определяет задачи, которые могут быть выполнены, в том числе следующие: звонки через голосовую связь по IP-протоколу, действия с сообщениями, платежи, обработку фотографий, тренировки, бронирование такси, CarPlay (только для производителей автомобилей), предварительный заказ столиков в ресторане (требуется дополнительная поддержка от Apple).

Существует два типа App Extensions (сервисов), которые используются для взаимодействия с Siri:

1) Intents Extension – обеспечивает взаимодействие Siri и Maps с контентом приложений и выполняет задачи, необходимые для реализации любого из поддерживаемых intents.

2) Intents UI Extension – предоставляет пользовательский интерфейс, который отображает внутренний контент либо Siri, либо Maps.

Очень важно, чтобы приложение обеспечивало Intents Extension для поддержки SiriKit. Однако можно и отказаться от предоставления UI Extension, так как Siri обрабатывает все взаимодействия с пользователем, используя стандартный, встроенный интерфейс. Intent Extension может поддерживать один или несколько Intents. Apple требует, чтобы мы, как разработчики, старались ограничивать число

Intent Extensions, так как это помогает поддерживать использование памяти на низком уровне и повышает производительность Siri.

Системы распознавания речи классифицируются:

- 1) по размеру словаря (ограниченный набор слов, словарь большого размера);
- 2) по зависимости от диктора (дикторозависимые и дикторонезависимые системы);
- 3) по типу речи (слитная или отдельная речь);
- 4) по назначению (системы диктовки, командные системы);
- 5) по используемому алгоритму (нейронные сети, скрытые Марковские модели, динамическое программирование);
- 6) по типу структурной единицы (фразы, слова, фонемы, дифоны, аллофоны);
- 7) по принципу выделения структурных единиц (распознавание по шаблону, выделение лексических элементов).

Помехозащищённость систем автоматического распознавания речи, обеспечивается, прежде всего, использованием двух механизмов:

- 1) использование нескольких, параллельно работающих, способов выделения одних и тех же элементов речевого сигнала на базе анализа акустического сигнала;
- 2) параллельное независимое использование сегментного (фонемного) и целостного восприятия слов в потоке речи.

Работа с речью полностью завязана на доступе к Интернету, и это неудивительно, так как используется тот же самый движок Siri. Разработчику даётся возможность либо распознавать живую речь, либо речь, записанную в файл.

Но так как Speech.framework не содержит специального API для работы с микрофоном, разработчику потребуется задействовать AV Audio Engine. AV Audio Engine – это пространство для узлов, таких как схема или граф. Из возможных типов узлов строится цепочка обработки сигнала. Соединяя их различными способами, мы получаем требуемый эффект. Рассмотрим работу с real-time диктовкой. Вся работа представлена на принципиальной схеме (рис. 1):

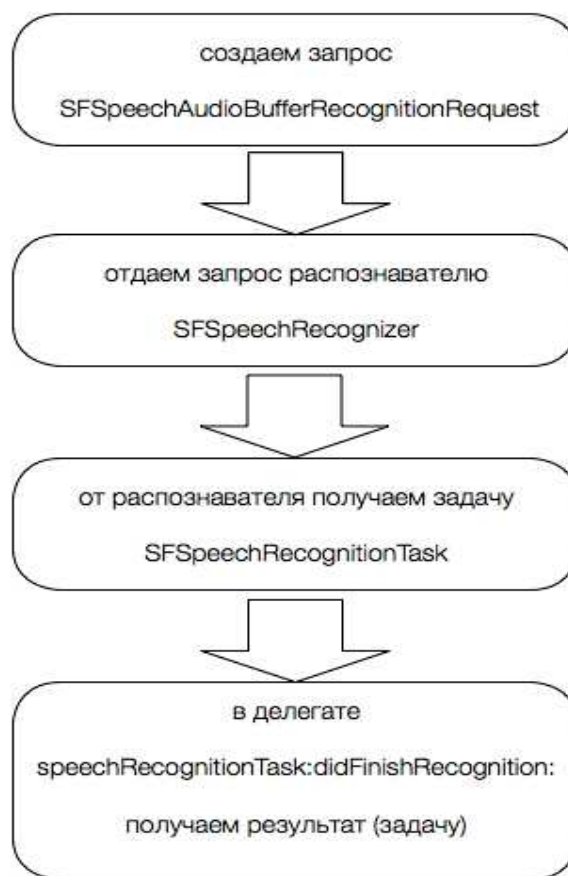


Рисунок 1

Преимущества преобразования голоса в текст:

- 1) возможность использовать свой голос, чтобы отправлять SMS-сообщения или электронные письма. Можно забыть о неудобном печатании и практически полностью ограничить использование клавиатуры;
- 2) возможность перевести текст на 32 иностранных языка;
- 3) корректировка текста с помощью голосовой диктовки или ввода с клавиатуры;
- 4) полная поддержка русского языка.

Независимо от того, насколько быстро мы вводим с клавиатуры текст, этот способ никогда не сможет работать быстрее, чем мы произносим слова. К счастью, современные технологии не стоят на месте, и преобразование голоса в текст затрагивает все больше и больше областей в нашей жизни. И скоро мы станем свидетелями серьезной конкурентной борьбы в области продвижения технологий преобразования голоса в текст. Чей алгоритм окажется лучше – покажет время.