СОДЕРЖАНИЕ

|  |  |
| --- | --- |
| Задание | 2 |
| Использованное ПО | 3 |
| Голосовые движки | 4 |
| Тестовый пример | 6 |
| Заключение | 7 |

Задание

Найти и проанализировать движки для синтеза русской речи.

Входные данные: текст на русском языке

Выходные данные: воспроизведение синтезированной речи в виде звукозаписи

Использованное ПО

Для анализа синтеза речи были использованы шесть движков: **Alyona**, **Nicolai**([http://www.acapela-group.com](http://www.acapela-group.com/)), **eSpeak** (<http://ru.wikipedia.org/wiki/ESpeak>), **Festival** (<http://ru.wikipedia.org/wiki/Festival>), **Igor**, **Google Tranclate** (<https://translate.google.ru/>).

Для ввода входных данных и прослушивания, настройки скорости воспроизведения и тембра была использована читалка “**Балаболка**” (<http://www.cross-plus-a.ru/balabolka.html>).

Так как **Google Tranclate** использовалась через интернет и не поддерживает сохранения звукозаписи на компьютер, была использована программа **Total Recorder** (<http://www.totalrecorder.com/>). Эта программа распараллеливает звуковой поток, приходящий на звуковую карту, и записывает его. Так как тривиальная версия Total Recorder` а сохраняет звуковой поток с карты в формате “.mp3”, чтобы привести его к общему формату(“.wav”), использовался **AIMP3 Audio Converter (**идет вместе с AIMP <http://aimp.ru/>).

Голосовые движки

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Хар-ки  Движки | Соблюдение знаков препинания | Качество речи | Особенности | Раскрытие сокращений | Правильность ударения | Пол диктора | Разработчик |
| Alyona | + | среднее | ударение ставится на первый слог (хитрость – выставление пробела перед ударным слогом) | частичное | - | Ж | **Acapela** |
| Nicolai | + | высокое | слова откладываются в голове, хорошее качество синтеза (лучшее из всех рассмотренных в табличке) | частичное | + | М | **Acapela** |
| eSpeak | + | низкое | очень низкое качество речи | нет | + | Робот | open source |
| Festival | + | выше среднего | самый сложный движок в установке | частичное | + | М/Ж | **Алан В. Блэк** из Исследовательского Центра речевых Технологий |
| Igor | - | низкое | движок отличается украинским говором (вообще это украинский движок, но русскую речь синтезирует вполне нормально) | нет | + | М | ukrvox |
| Google Tranclate | + | выше среднего | голос немного нервирует, присутствуют скачки голоса (будто диктор иногда проявляет эмоции) | частичное | + | Ж | Google Inc. |

К сожалению, два движка не удалось установить: Loquendo (Olga) и ScanSoft RealSpeak (Katerina).

К сожалению, английская речь синтезируется гораздо лучше русской (не рассмотренными русскоголосыми движками, а иностранными (тот же eSpeek, Microsoft Mike, Mary)).

Также есть движки для эмоционально окрашенной речи (для английского и японского языка), но рассмотрены не были.

Тестовый пример

Тестовым примером служит вольный перевод обращения [Дж. Картера](http://ru.wikipedia.org/wiki/%D0%9A%D0%B0%D1%80%D1%82%D0%B5%D1%80,_%D0%94%D0%B6%D0%B8%D0%BC%D0%BC%D0%B8), записанное на золотой пластинке «Вояджера» (<http://ru.wikipedia.org/wiki/%D0%97%D0%BE%D0%BB%D0%BE%D1%82%D0%B0%D1%8F_%D0%BF%D0%BB%D0%B0%D1%81%D1%82%D0%B8%D0%BD%D0%BA%D0%B0_%C2%AB%D0%92%D0%BE%D1%8F%D0%B4%D0%B6%D0%B5%D1%80%D0%B0%C2%BB>).

Тестовый пример был выбран за наличие сокращения и знаков пунктуации.

Текст:

*Этот аппарат создан в США, стране с населением 240 млн. человек среди 4-миллиардного населения Земли. Человечество всё ещё разделено на отдельные нации и государства, но страны быстро идут к единой земной цивилизации.  
Мы направляем в космос это послание. Оно, вероятно, выживет в течение миллиарда лет нашего будущего, когда наша цивилизация изменится и полностью изменит лик Земли… Если какая-либо цивилизация перехватит „Вояджер“ и сможет понять смысл этого диска — вот наше послание:  
Это — подарок от маленького далёкого мира: наши звуки, наша наука, наши изображения, наша музыка, наши мысли и чувства. Мы пытаемся выжить в наше время, чтобы жить и в вашем. Мы надеемся, настанет день, когда будут решены проблемы, перед которыми мы стоим сегодня, и мы присоединимся к галактической цивилизации. Эти записи представляют наши надежды, нашу решимость и нашу добрую волю в этой Вселенной, огромной и внушающей благоговение.*

В архиве с протоколом находятся пять файлов в формате “.wav”, содержащие звукозаписи выходных данных каждого из рассмотренных движков (кроме звукозаписи Festival` я, с ним в этом плане возникли большие проблемы).

Заключение

Самым лучшим движком из рассмотренных на мой взгляд является движок от **Acapel` ы,** Nicolai. Приятный голос, хорошая постановка ударений, работы с пунктуацией, нахождение большинства сокращений и хорошо воспринимаемая речь дает этому движку большое преимущество перед остальными движками.

Самым худшим является движок eSpeak за его «машинный» голос (многим людям более привычен людской голос, хотя мне больше нравится машинный, но против природы не попрешь), низкое качество речи и никакого обращения внимания на сокращения.

Голосовые движки очень интересны. Их можно выгодно внедрить в производство, в информационно – справочные системы, объявление о следующей станции в электричке, объявление время отправления поезда, различные массовые оповещения и конечно же в дальнейшем в музыке и т.д.