



DSpeech v1.56.2 (от Dimio)

Обновленная версия находится на домашней странице: <http://dimio.altervista.org/>

Описание

«**DSpeech**» – программа преобразования текста в речь (TTS) со встроенной возможностью автоматического распознавания речи (ASR). Она способна читать текст и выбирать предложения для произнесения в соответствии с голосовыми командами пользователя. Создана для быстрой работы и располагает ценными практическими функциями, необходимыми для программ такого типа. В то же время является ненавязчивой и нетребовательной к ресурсам: не требует установки, не встраивается в систему, имеет малый размер, запускается мгновенно и не делает записей в реестре.

Основные особенности «DSpeech»

- ✓ Позволяет сохранять выходной файл в файлах WAV или MP3.
- ✓ Позволяет быстро выбирать различные голоса и соединять их между собой для создания диалогов.
- ✓ Полноценная система распознавания речи, которая за счет простых команд позволяет создавать диалоги взаимодействия с пользователем.
- ✓ Позволяет изменять отдельные голоса отдельно.
- ✓ Позволяет с помощью особых тегов динамически изменять характеристики голосов во время воспроизведения (скорость, громкость и частота), вставлять паузы, выделять голосом понятия и читать по буквам.
- ✓ Позволяет доставать и автоматически воспроизводить содержание буфера обмена.
- ✓ Поддерживает все речевые пакеты, совместимые с SAPI4 и SAPI5.

В числе дополнительных возможностей

- ✓ Во время запуска позволяет автоматически загружать последний открытый файл с сохранением позиции чтения.
- ✓ Поддерживает командную строку и может быть использована для создания аудиокниг даже без графической оболочки.
- ✓ Позволяет указать формат выходного аудиофайла, что может оказаться полезным, например, в ситуациях, когда требуется файл WAV с определенными характеристиками.
- ✓ Позволяет создавать вставки из файлов MP3 и WAV, используя особые команды. Это может быть использовано для внесения в чтение особых эффектов в виде приступа кашля, смеха либо вашего музыкального файла.
- ✓ При сохранении в MP3 можно выбирать его качество с учетом либо выходного размера, либо качества звучания.
- ✓ Теперь можно преобразовывать тексты в MP3 с разбиением на файлы по 5, 10, 15 минут.
- ✓ В файл CustomTAG.TXT можно ввести некоторые личные теги, которые после появятся в контекстном меню (например, теги выразительного чтения для пакета «Loquendo»).

Преобразование текста в речь

Посредством контекстного меню (правая кнопка мыши) можно указать, каким голосом произносить предложение, что делает возможным создание диалогов из различных голосов.

Аналогично, возможна вставка особых тегов, позволяющих изменять характеристики голоса во время чтения (скорость, громкость, частота).

Настройки звука

Можно выбрать устройство вывода звука и качество аудиопотока для выбранного голоса. Рекомендуется всегда использовать одинаковое качество потока (бит/сек) для всех используемых голосов, иначе возможны потери качества звука и «эффект консервной банки».

Большинство голосов рассчитаны на 16-битный монозвук с частотой 16 КГц. Голоса «Майкрософта» используют 16-битный монозвук с частотой 22 КГц. Эти настройки особенно важны во время преобразования в файлы WAV или MP3.

Преобразование в файлы WAV или MP3

«*Dspeech*» может преобразовывать текст в файл WAV или MP3. При выборе MP3 предоставляется возможность указать степень сжатия. Можно выбрать максимальный размер в самом лучшем качестве, минимальный размер с низким качеством, а также сбалансированный вариант.

Редактирование текста

Помимо основных функций правки текста (найти, заменить, вырезать и пр.) присутствует особая функция – удаление лишних возвратов каретки. Она используется для удаления всех существующих в тексте лишних возвратов каретки, которые могут повлиять на плавность чтения искусственными голосами. В результате постраничной верстки текст может содержать возвраты каретки, которые ухудшат качество чтения.

Создание аудиокниги

Для создания аудиокниг присутствуют несколько настроек; в частности, возможность разбивать текст на файлы по несколько минут каждый, включая разбиение по меткам в тексте в указанных местах. В таком случае необходимо вставлять команду #BREAK каждый раз там, где необходимо разбиение.

Используя команду меню «Добавить...», можно объединить несколько текстовых файлов, вставляя их один за другим. Между двумя открытыми таким образом файлами будет автоматически вставлена команда #BREAK, что предусматривает разделение отдельных текстов при записи. Можно добавить #BREAK в начало текста для улучшения совместимости с проигрывателями CD и старыми MP3-плеерами.

Распознавание речи

«*Dspeech*» поддерживает систему распознавания речи, которая позволяет создавать диалоги взаимодействия с помощью простых команд.

Пример

Пользователь: «Компьютер!»
ПК: «Включен. Кто вы?»
Пользователь: «Димио».
ПК: «Приветствую».
и т. д.

Система команд схожа с языком BASIC. На данный момент поддерживаются следующие команды:

```
#VOICE Название голоса
#I GIVE
#EXIT DO
#LOOP
#RECOGNIZE Слово1, [Слово2], [OTHER_WORDS]...
#RECOGNIZE_WITH_TIMEOUT Секунд, Слово1, [Слово2], [OTHER_WORDS]...
#IF RECOGNIZED Слово1, [Слово2], [OTHER_WORDS]...
#IF TIMEOUT
#END IF
#CALL название
#Sub название
#END SUB
#RANDOM
#HOUSES
#END RANDOM
#EXECUTE путь к файлу программы
#OPEN FileToSpeech.txt
#STOP
#BREAK
#PLAY FileName.wav
#WAIT количество секунд
#CLOSE
```

Все эти команды можно найти в контекстном меню (правая кнопка мыши) с примерами. Подробный синтаксис здесь описан не будет – его можно понять из примеров. Примером команд для голосового распознавания могут быть следующие:

```
#VOICE Marco
Я Ангел – компьютер Дмитрия. Кто вы?
#DO
#RECOGNIZE Дмитрий, Глория, OTHER_WORDS
#IF RECOGNIZED Дмитрий
    Ангел ожидает распоряжений.
#EXIT DO
#END IF
#IF RECOGNIZED Глория
    Вы ошиблись компьютером. Ваш тот, что рядом.
#EXIT DO
#END IF
#IF RECOGNIZED OTHER_WORDS
#RANDOM
#CASE
    Не могли бы вы повторить? Я не расслышал имени.
#CASE
    Что вы сказали? Не могли бы вы повторить?
#CASE
    Я не понял, что вы сказали. Возможно, ваше имя мне не известно.
#END RANDOM
```

#END IF
#LOOP

К сожалению, на данный момент система распознавания основана на фонетике английского языка, ввиду чего может потребоваться особое написание слов. Например, «Russia» будет распознано, если его написать «Russya».

Сочетания клавиш

Все функции программы могут быть вызваны сочетаниями клавиш.

Поддерживаются следующие сочетания:

F1 = переход в окно редактирования текста
F4 = пауза/воспроизведение
F5 = читать/не читать
F6 или ALT + UP = прочитать предыдущую строку
F7 или ALT + LEFT = прочитать текущую строку
F8 или ALT + DOWN = прочитать следующую строку
F9 или ALT + RIGHT = прочитать от курсора
F11 = переход к предыдущему голосу
F12 = переход к следующему голосу
ESC = стоп
ALT+1 = увеличить громкость
ALT+2 = уменьшить громкость
ALT+3 = увеличить скорость
ALT+4 = уменьшить скорость
ALT+5 = повысить тон
ALT+6 = понизить тон

Командная строка

Возможно указать название файла, который будет открыт автоматически и воспроизведен. Это позволяет автоматически выполнять команду в виде скрипта.

Синтаксис

DSPEECH.ExE [/Play] [/Speak] [/Wav] [/Mp3] [/Ogg] [/Hidden|/HiddenFix] [FileToSpeech.txt]

Примеры командной строки:

- Для открытия файла:

DSpeech.exe source.txt

- Чтобы начать воспроизведение файла:

DSpeech.exe /Play source.txt

- Чтобы вслух произнести фразу:

DSpeech.exe /Speak Hello!

- Чтобы преобразовать текстовый файл в mp3:

DSpeech.exe /mp3 source.txt [destination.mp3]

- Чтобы преобразовать текстовый файл в ogg:

DSpeech.exe /ogg source.txt [destination.ogg]

- Чтобы преобразовать текстовый файл в wav:

DSpeech.exe /wav source.txt [destination.wav]

Системные требования

Для «**Windows NT/2000**» требуется установка пакета MS-SAPI5.1. Он есть на сайте «Майкрософта», а также здесь:

<http://aldostools.mysite4now.com/sapi51.msi>

<http://www.arlington.com.au/sapi51.msi>

Если требуется распознавание речи на «**Windows NT/2000**», необходимо установить средство распознавания речи, для чего скачать с сайта «Майкрософта» весь пакет SAPI 5.1 SDKs (60 MB) либо отдельно средство распознавания здесь:

<http://clans.gameclubcentral.com/shoot/SR.zip>

«Windows XP/2003/VISTA»

Установка дополнительных пакетов не требуется.

Минимальное разрешение экрана для «**Dspeech**» – 1024x768.

На системах семейства «**Windows 9x**» программа не работает.

Голоса

«**Dspeech**» использует установленные в системе голоса. На «Windows XP» есть только «Microsoft SAM» – английский. Если на «Windows NT/2000» установлен пакет MS-SAPI5.1, будет доступно два других голоса – Майк и Мэри, также английских. Пользователи «XP» могут скачать их отсюда:

<http://download.microsoft.com/download/speechSDK/SDK/5.1/WXP/EN-US/Sp5TTIntXP.exe>

Указанные голоса лучше, чем SAM, но и их качество, по сравнению с разработками сторонних производителей, оставляет желать лучшего. Поэтому рекомендуется установка голосовых

пакетов сторонних производителей. К сожалению, лучшие продукты такого рода являются платными, и даже их приобретение не гарантирует беспроблемной установки. В любом случае, среди них следует отметить такие:

[итальянские]

Asapela (чистая и связная речь, но не лучший вариант в плане естественности).

Serstral (с хорошим соотношением цены/качества; качество уступает самым дорогим голосовым пакетам, но все же годится для использования).

Loquendo (вероятно, наилучший вариант, особенно в плане естественности и выразительности; цена не малая).

RealSpeak (весьма хорош).

[итальянские]

VoiceWare (хорошие голоса; итальянский не включен).

Замечания

При вставке большого MP3-файла в текст может возникнуть небольшая пауза между чтением предыдущего текста и воспроизведением аудиофайла. Это нормально и не является сбоем. При сохранении результата в файлах WAV или MP3 пауза полностью исчезает.

Для сжатия в MP3 используется кодек «Blades» (www.mp3dev.org). Он соответствует файлу «Lame.exe», включенному в пакет.

Программист:

Dimitrios Coutsoumbas (Dimio)

SKYPE : katafratto

ICQ : 145633952

Почта : cyberdimio@gmail.com

Страница : <http://dimio.altervista.org/>

Испытатель бета-версии:

Talksina (talksina@gmail.com)