

КОМПЬЮТЕРНЫЕ РАЗРАБОТКИ ДЛЯ ДЕТЕЙ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ

Д.В.Копейка, ст. гр.ЭНМ09
Л.В.Славинская, ст. препод. каф. ВМиП

Донецкий национальный технический университет

Child's limited facilities is not only medical diagnosis with a long treatment, but it is also the search of the place in society, necessity of getting education. Traditional stading technology support the necessary standard only for children going to the general institutions. In this paper we consider the modern speech technologies.

Ограниченные возможности ребенка - это не столько медицинский диагноз, предполагающий длительное лечение, но это еще и поиск своего места в обществе, необходимость получения образования. Традиционные технологии обучения обеспечивают необходимый уровень образования лишь для детей, посещающих общеобразовательные учреждения. В данной работе рассматриваются современные речевые технологии.

Обмежені можливості дитяти - це не стільки медичний діагноз, який передбачає тривале лікування, але це ще і пошук свого місця в суспільстві, необхідність здобування освіти. Традиційні технології учення забезпечують необхідний рівень освіти лише для дітей, відвідуючих загальноосвітні установи. У даній роботі розглядається сучасні мовні технології.

РЕЧЕВЫЕ ТЕХНОЛОГИИ РАСПОЗНОВАНИЕ КОМАНД РЕЧЕВОЙ СЕРВЕР

Дети-инвалиды, а также больные дети, которые не могут посещать общеобразовательные учреждения в силу ограничений по состоянию здоровья, лишены многих возможностей в получении качественного образования. Проблема помощи таким детям является одной из важнейших в ряду множества проблем, стоящих перед отечественным образованием.

Получение знаний обычным путем для больных и инвалидов ограничено или невозможно. Расширить доступ инвалидов к образовательным программам и сильной трудовой деятельности можно, широко внедряя новые информационно-образовательные технологии.

В данной работе рассматривается информационно-образовательные технологии лишь одного направления – *современные речевые технологии*.

Это готовые программы и системы широкого класса: программы голосового управления компьютером, домашним и офисным оборудованием, мобиль-

ным телефоном, системы голосового разграничения доступа и другие системы, в которых применяются современные речевые технологии.

Синтез и распознавание речи

Существует два подхода к распознаванию речи. Первый подход реализует распознавание элементов речи по образцу и применяется в различного рода системах голосового управления. Вторым подходом основан на выделении в речи лексических элементов — фонем, аллофонов, морфем и т.д. Этот подход пригоден для создания систем диктовки текста. Можно выделить системы распознавания речи, требующие обучения и зависящие от диктора, а также системы, способные работать без предварительного обучения и, следовательно, не зависящие от диктора.

Перед тем как приступить к выделению из речи лексических элементов, необходимо выполнить предварительную обработку речевого сигнала. В ходе этой обработки из сигнала удаляются шумы, выполняется частотная фильтрация и оцифровка, а также нормализация уровня сигнала.

Рассмотрим две методики выделения из речи лексических элементов.

Первая методика предполагает использование дискретного преобразования Фурье. Непосредственно лексические элементы выделяются из оцифрованной речи при помощи нейронной сети, способной к обучению. При этом речь представляется в виде некоторого набора числовых параметров, так как нейронные сети работают именно с наборами таких параметров.

Вторая методика выделения лексических элементов речи основана на применении вейвлет-преобразований. В отличие от дискретного преобразования Фурье, этот метод исключает потерю информации о временных характеристиках обрабатываемых сигналов. При использовании вейвлет-преобразований входной сигнал раскладывается не в базисе периодических функций (как в дискретном преобразовании Фурье), а в базисе функций, характеризующих как частоту, так и время [1].

Системы голосового управления

Все системы голосового управления компьютером и другим оборудованием можно условно разделить на два класса.

К первому классу относятся устройства, назначением которых является облегчение работы с компьютером. Такие устройства используются в качестве дополнительного канала управления наряду с основными и традиционными каналами — клавиатура и мышь, а также экран компьютера (который служит каналом обратной связи). В устройствах второго класса речевой интерфейс является основным и единственным. Эти устройства доступны тем, кто не может работать с клавиатурой и экраном компьютера либо из-за проблем со здоровьем. Рассмотрим несколько систем голосового управления, созданных в компании «Центр речевых технологий» [2]. Это программы VoiceNavigator и Truffaldino, библиотека распознавания голосовых команд VoiceKeyKit. Все они, относятся к первому классу, т.к. для работы с ними необходимы традиционные средства ввода и вывода информации.

Программа VoiceNavigator

Программа VoiceNavigator (рис.1) является типичным представителем программ голосового управления компьютером. Она позволяет пользователю запускать приложения голосом, не дотрагиваясь до клавиатуры, и выполнять произвольно заданные команды.

Рис.1. Программа VoiceNavigator

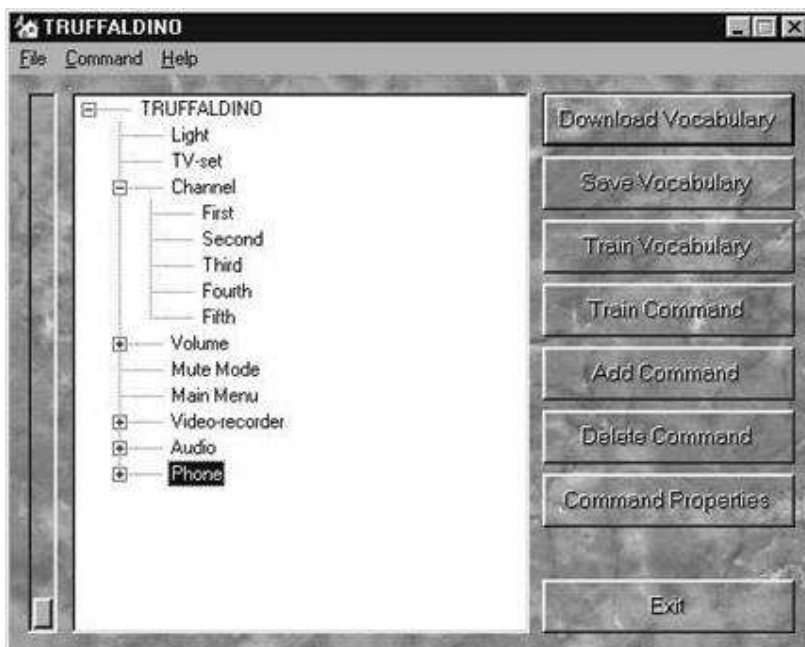


Перед использованием программы VoiceNavigator ее необходимо обучить, произнеся в микрофон слова команд. Так как программа VoiceNavigator распознает команды по образцам, то команды можно произносить на любом языке и любым голосом. Чтобы программа начала распознавать голосовые команды, ее необходимо «разбудить», произнеся ключевое слово. После этого программа будет реагировать только на Ваши команды, игнорируя остальные звуки. В программе есть функция голосового ответа-подтверждения команд. Эта функция позволяет убедиться, что Ваша команда распознана системой и готова для выполнения. Программа VoiceNavigator нетребовательна к ресурсам компьютера. Ее можно использовать в компьютере, оборудованном процессором с тактовой частотой 200 МГц или выше, причем для ввода звуковых команд подойдет любой звуковой адаптер, например, Creative Sound Blaster.

Программа Truffaldino

Программа Truffaldino (рис.2) предназначена для управления домашним или офисным оборудованием, таким как телевизор, домофон, видеоманитофон, аудио-система, кондиционер, микроволновая печь, система освещения, телефон и т.п.

Рис.2. Программа Truffaldino



Эта программа пригодится людям с ограниченными физическими возможностями, не способными управляться с домашней техникой обычным способом. После предварительного обучения программа Truffaldino способна распо-

распознавать несколько десятков команд, произнесенных на любом языке и любым голосом. Последнее возможно благодаря тому, что команды распознаются по голосовым шаблонам. Особенностью программы является хорошая работа в условиях шумов. Помимо ключевой фразы, предназначенной для «пробуждения» программы Truffaldino, имеется функция ограничения доступа. С помощью этой функции можно заблокировать запуск программы посторонними людьми.

В программе Truffaldino есть функция голосового ответа-подтверждения команд, аналогичная такой же функции программы VoiceNavigator. Эта функция позволяет убедиться, что Ваша команда распознана системой и готова для выполнения. Системные требования программы Truffaldino аналогичны требованиям программы VoiceNavigator. Однако для управления оборудованием необходимо применение дополнительного устройства, такого как Infra-Red X-tender.

Библиотека распознавания голосовых команд VoiceCom

Библиотека распознавания голосовых команд VoiceCom составляет ядро описанных выше программ VoiceNavigator и Truffaldino. С ее помощью разработчики могут добавить голосовое управление в создаваемые ими приложения.

Разработка собственных средств голосового управления требует специальных знаний. Она может отнять у разработчиков приложений немало времени и сил. Однако, воспользовавшись готовой библиотекой VoiceCom, разработчики могут легко добавить в приложения следующие функциональные возможности:

- управление оборудованием с помощью голоса;
- выполнение речевых запросов к базам данных через микрофон или даже по телефону;
- поиск по ключевым словам в звуковых WAV-файлах

Следует отметить, что библиотека VoiceCom позволяет встраивать голосовые функции не только в обычные программы для персональных компьютеров, но и в автономные устройства, оборудованные цифровыми сигнальными процессорами DSP.

Алгоритмы, реализованные в библиотеке распознавания голосовых команд VoiceCom, обладают высоким быстродействием, нетребовательны к объему оперативной памяти и способны адаптироваться к шумам.

Библиотека VoiceCom обеспечивает распознавание команд, произнесенных любым голосом и на любом языке. При этом имеется возможность структурирования для практически неограниченного словаря.

При этом алгоритмы позволяют распознавать 100-200 команд с предварительным обучением для каждого диктора, и 30-50 команд для любого диктора (в режиме, не зависящем от диктора). Если команды произносятся по телефону, то алгоритмы библиотеки VoiceCom позволяют распознать 10-20 слов, произнесенных любым диктором.

В библиотеке также реализована возможность активации распознавания команд по ключевому слову, исключая неожиданные реакции системы на посторонние звуки.

Программы для диктовки текста

Распознавание речи чаще всего используется для выполнения голосовых команд. Однако сегодня существуют мощные приложения, способные не только распознавать и выполнять команды, но и распознавать речь в режиме диктовки. Как правило, такие приложения либо снабжаются собственным редактором текста, либо способны работать с любыми редакторами текста и таблиц, например, такими, как Microsoft Word и Microsoft Excel.

Кроме того, система распознавания речи (увы, не способная работать с русским языком), входит в комплект современного офисного пакета Microsoft Office XP.

Программа DragonDictate

Примером системы распознавания речи, способной работать как в режиме команд, так и в режиме диктовки текста может служить программа DragonDictate. Она получила свое развитие в виде линейки программ Dragon NaturallySpeaking.

После установки и настройки программы DragonDictate и Dragon NaturallySpeaking позволяют запускать при помощи голоса приложения Windows, управлять работой этих приложений в командном режиме, а также диктовать текст. К сожалению, программы DragonDictate и Dragon NaturallySpeaking не работают с русским языком.

Основные приемы работы с программой DragonDictate

Включение и выключение микрофона

Если микрофон постоянно включен, система распознавания речи может реагировать на посторонние шумы различного рода. В зависимости от того, в каком режиме работает программа (в командном режиме или в режиме диктовки) эти шумы могут приводить либо к выполнению ложных команд, либо к появлению «мусора» в редактируемом документе.

Микрофон можно включать и выключать с помощью кнопки с изображением микрофона, щелкая ее левой клавишей мыши, либо с помощью клавиши +, расположенной на дополнительной клавиатуре.

Для голосового управления микрофоном следует использовать команды Wake Up и Go to Sleep. Первая из этих команд включает микрофон, а вторая — выключает. Распознанная и выполненная команда отображается в правой части инструментальной голосовой панели и заключена в квадратные скобки. На рис. 3 показан результат выполнения команды Go to Sleep, а на рис. 4 — результат выполнения команды Wake Up.



Рис.3. Результат выполнения команды Go to Sleep



4 — результат выполнения команды Wake Up.

Распознавание речи в Microsoft Office XP

Осознавая актуальность технологий распознавания и синтеза речи, компания Microsoft встроила в свои операционные системы Microsoft Windows 2000 и Microsoft Windows XP специальный речевой программный интерфейс Speech API. Следует, однако, заметить, что наличие этого интерфейса не означает, что перечисленные выше операционные системы способны распознавать речь или говорить. Интерфейс только позволяет разработчикам систем распознавания и синтеза речи встраивать их в операционную систему стандартным способом. В составе американского, японского и упрощенного китайского пакета офисных приложений Microsoft Office XP поставляются средства распознавания и синтеза речи. Что же касается русской версии Microsoft Office XP, то, к большому сожалению, там эти средства отсутствуют.

Говорящий Linux.

Речевой сервер Speech-Dispatcher версии 0.7

Речевой сервер с открытым исходным кодом Speech-Dispatcher версии 0.7 предназначен для работы в дистрибутивах GNU/Linux. В новой версии по умолчанию используются сокеты Unix, что позволяет избежать проблем с выбором номера порта и решает некоторые вопросы безопасности; поддерживается автозапуск сервера по требованию пользователя; переработан вывод звука через pulse audio; сервер по умолчанию запускается как пользовательская служба, а не как системная; предусмотрен механизм отката к использованию alsa, если pulse audio оказывается недоступным; обновлена документация; сделаны другие исправления и улучшения.

Vinux 3.0

Vinux 3.0 специально адаптирован для нужд незрячих и слабовидящих пользователей. Этот выпуск Vinux включает три программы чтения экрана и две программы увеличения экрана, поддерживает брайлевские дисплеи, подключаемые через USB, позволяет изменять цветовую схему и шрифты. Vinux доступен в двух вариантах: как установочный Live CD и в виде deb-пакета, который автоматически преобразует установленный вариант Ubuntu 10.4 в систему Vinux, более доступную для незрячих и слабовидящих пользователей.

Литература

1. Д. Х. Клэнтт, Дж. А. Барнет, М. И. Бернстейн и др. Методы автоматического распознавания речи: В 2-х книгах. Пер. с англ. /Под ред. У. Ли. – М.: Мир, 1983
1. А.В. Фролов, Г.В. Фролов. Синтез и распознавание речи. Современные решения – М.: Мир, 2003