
ВСПОМОГАТЕЛЬНАЯ СИСТЕМА РЕЧЕВОГО УПРАВЛЕНИЯ АВТОМОБИЛЕМ НА ОСНОВЕ ТЕХНОЛОГИИ SPEECH API

А.И. Дивеев

Вычислительный центр им. А.А. Дородницына РАН
ул. Вавилова, 40, Москва, Россия, 119991

Атиенсия Хосе

Кафедра Кибернетики и мехатроники
Российский университет дружбы народов
ул. Миклухо-Маклая, 6, Москва, Россия, 117198

В статье рассматривается технология разработки дополнительной системы речевого управления автомобилем на основе пакета API функции Windows Speech. Приведены результаты экспериментального исследования реализации системы. Настоящая статья посвящена созданию и исследованию алгоритмов речевой диалоговой системы управления.

Рассматривается система управления подвижным объектом. В контуре управления объектом присутствует человек-оператор, который осуществляет обработку поступающих с приборов данных и на основе полученной информации осуществляет основное управление движущимся объектом. Помимо основного контура управления в системе могут быть дополнительные контуры, также реализующие свое воздействие на объект и оказывающие влияние на качество управления с помощью человека-оператора.

Одним из дополнительных контуров управления подвижным объектом, замыкаемым через человека-оператора, является управление с помощью речевых команд. Человек-оператор воспроизводит дополнительные команды, помимо используемых им управляющих воздействий, в виде отдельно звучащих слов. Блок управления, реализованный в виде микропроцессорной системы, распознает поступающие команды и осуществляет дополнительные операции, повышающие качество управления или обеспечивающие изменения режимов управления, например переход с автоматического режима управления на ручной режим.

Управление автомобилем осуществляется на основе полученной информации с помощью воздействия на автомобиль через рулевую колонку, педали газа и тормоза. Помимо основных видов управления, которые должны быть реализованы в режиме реального времени, водитель автомобиля осуществляет также управление и по дополнительным контурам, в которых режим реального времени не может строго не соблюдаться. Наиболее часто встречающееся в автомобиле управление по дополнительному контуру — это переключение скорости передачи с помощью рычага переключения и педали сцепления. Эта операция, хотя и является обязательной, но не требует строгого соблюдения временной реализации. В автомобиле также к дополнительным контурам управления следует отнести: включение дополнительного заднего или переднего привода, включения автомата блокировки тормозов, включения и выключения световых сигналов поворотных индикаторов, фар дальнего или ближнего света, включения дворников, печи обогрева салона, кондиционера и т.п.

Во многих современных автомобилях управление осуществляется по дополнительному контуру с помощью интеллектуальной системы управления и бортового компьютера. Компьютер сам производит включение режимов управления и решает за водителя необходимость их осуществления. Однако такая форма решения задачи дополнительного управления часто вызывает раздражение водителей из-за неверно принятых компьютером решений. Наиболее известный пример — это автоматическая коробка передач. Не все водители согласны с ее работой и не всегда ее работа соответствует наилучшему режиму движения.

Использование автоматических и интеллектуальных систем в дополнительных контурах управления автомобилем продиктовано необходимостью разгрузить водителя от отвлекающих его механических действий. Как бы ни было просто переключать передачи, все равно это действие требует временных затрат и ослабление контроля в основном контуре. Автоматизация данных действий сегодня не может быть реализована в достаточной мере.

Для решения данной проблемы может быть предложен другой способ — осуществление управления по дополнительному контуру через речевые команды. Водитель, не отвлекаясь от основного контура управления, посылает в систему речевого управления дополнительную команду и осуществляет таким образом нужные переключения режимов управления.

Структурная схема системы управления подвижным объектом с дополнительным контуром управления по речевой команде представлена на рис. 1. На рисунке показаны два контура управления — основной и дополнительный. В контуре дополнительного управления размещен блок речевого управления, который осуществляет реализацию всех необходимых действий по распознаванию речевой команды и выработки управляющего сигнала.

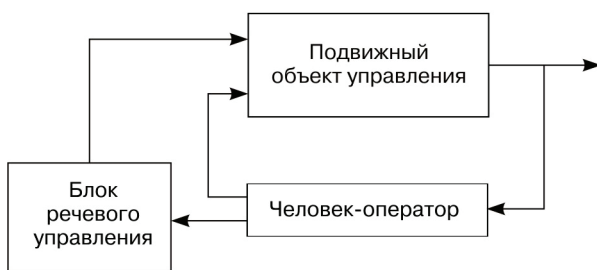


Рис. 1. Система речевого управления как дополнительный контур управления

Задача заключается в том, чтобы разработать и исследовать необходимые алгоритмы, которые реализуются в блоке речевого управления. Разрабатываемые алгоритмы должны осуществить следующие действия: анализ речевой команды; распознавание речевой команды; формирование внутреннего представления речевой команды; обработку информации во внутреннем представлении; формирование управляющего сигнала.

Рассматривается реализация системы распознавания речи, с помощью технологии Speech API в среде Windows. На рис. 2 приведена структурная схема системы распознавания.

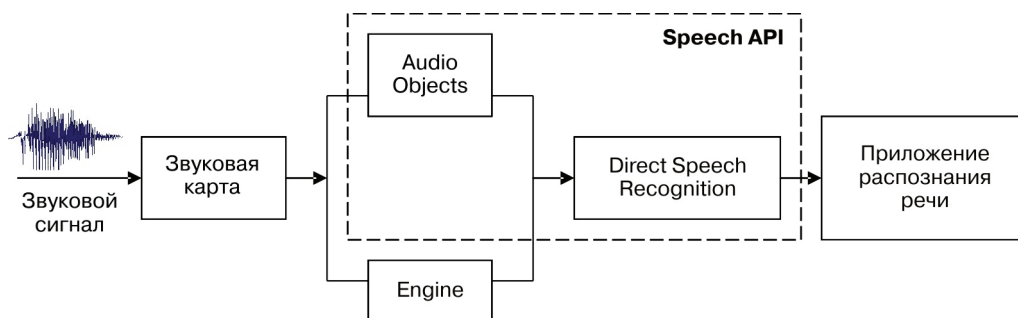


Рис. 2. Структурная схема системы распознавания речи

На рисунке блок *Audio Objects* обеспечивает доступ к звуковым устройствам; блок *Direct Speech Recognition* предназначен для прямого распознавания речи на низкоуровневом интерфейсе с помощью блока *Engine* построенного, как правило, на основе скрытых марковских моделей. Результаты экспериментов показали высокую эффективность системы и способность реагировать на большинство команд без обучения к голосу водителя.

Алгоритм распознавания речи с помощью *Direct Speech Recognition API* состоит из следующих этапов:

1) обращаемся к устройству, являющемуся источником аудиоданных, через выбранный интерфейс. В нашем случае это *Multimedia Audio Source*;

2) осуществляем поиск модулей распознавания речи по определенным критериям;

3) применяем выбор и получаем указатель на центральный интерфейс движка *ISRCentral*;

4) загружаем необходимую грамматику в объект *Grammar Compiler* и компилируем ее;

5) создаем объект, реализующий интерфейс уведомлений *ISRNotifySink* для получения информации о состоянии распознавания речи и регистрируем его через метод *Register* центрального интерфейса объекта *Engine*;

6) создаем объект, реализующий интерфейс *ISRGramNotifySink* для получения уведомлений о распознанных фразах, и загружаем скомпилированную грамматику в движок через объект *Grammar Compiler*, при этом передавая указатель на интерфейс уведомлений. Взамен получаем указатель на интерфейс *IUnknow* объекта *Grammar*;

7) через возвращенный указатель получаем указатель на интерфейс управления грамматикой *ISRGramCommon*;

8) через *ISRGramCommon* активизируем движок на распознавания речи по загруженной грамматике.

На рис. 3 показана программа распознавания изолированных слов, а в таблице представлены результаты экспериментальной проверки системы распознавания слов.

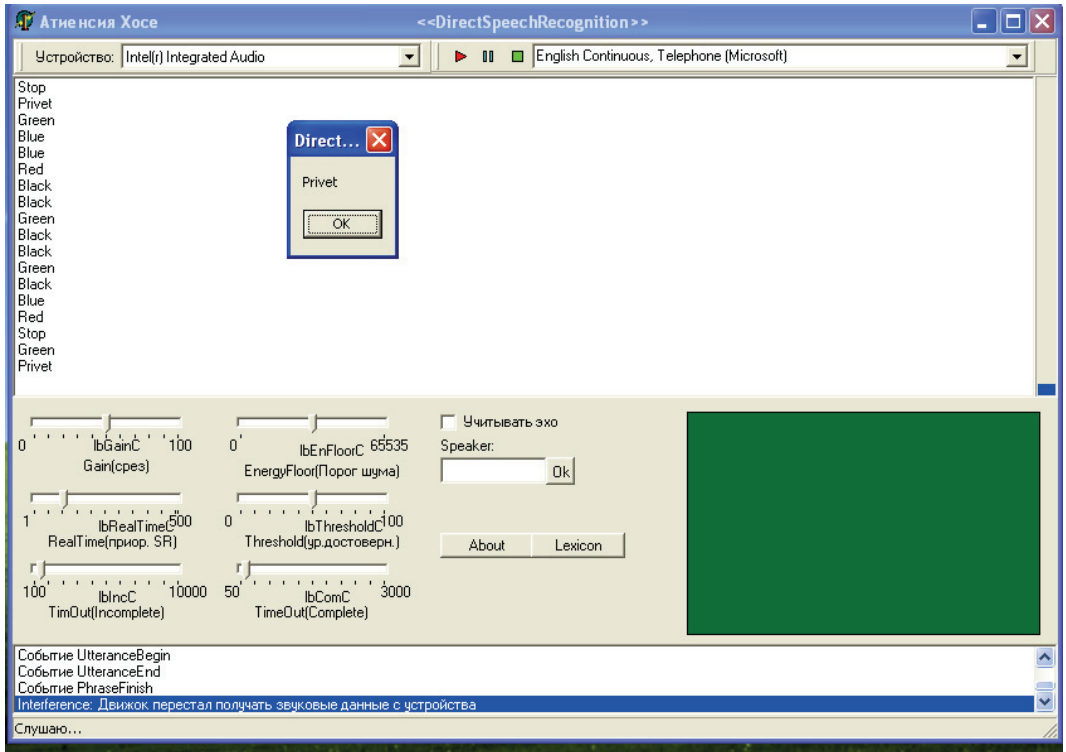


Рис. 3. Программа распознавания изолированных слов, в среде программирования Delphi 7.0

Таблица

**Экспериментальная проверка
системы распознавания изолированных слов**
Исходные данные

Диктор		Слова							Процент распознавания
Ф.И.О.	С/Б	Stop	Green	Blue	Red	White	Privet		
Дивеев А.И.	д.т.н.	1 1 1	1 1 1	1 1 1	0 1 1	1 1 1	1 1 1	94,44	
Софронова Е.А.	к.т.н.	1 1 1	1 1 1	1 0 1	1 1 0	1 1 1	1 1 1	88,89	
Васкес Мигель	2015223	1 1 1	1 1 1	1 1 1	1 1 1	1 1 1	1 1 1	94,44	
Осэс Пучков	2015154	1 1 1	1 1 1	1 1 1	0 1 0	1 1 1	1 1 1	88,89	
Браво Хавьер	1035190	1 1 1	1 1 1	1 1 1	1 1 1	1 1 1	1 1 1	100,0	
Ср. процент распознавания								93,33	

ЛИТЕРАТУРА

- [1] Рабинер Л.Р., Шафер Р.В. Цифровая обработка речевых сигналов / Пер. с англ.; под ред. М.В. Назарова и Ю.Н. Прохорова. — М.: Радио и связь, 1981.
- [2] Марпл С.Л. Цифровой спектральный анализ и его приложения. — М.: Мир, 1990.
- [3] Цифровая обработка сигналов. 2-е изд. — СПб: Питер.
- [4] Буторин Д.Н. MS Agent и Speech API в Delphi — СПб: БХВ-Петербург, 2005.

SUPPORTIVE SYSTEM OF SPEECH CONTROL OF A CAR BASED ON SPEECH API

A.I. Diveev

Dorodnicyn Computing Centre of the Russian Academy of Sciences
Vavilov str., 40, Moscow, Russia, 119333

Jose Atiensia

Cybernetic and mechatronics departament
Peoples' Friendship University of Russia
Miklukho-Maklaya str., 6, Moscow, Russia, 117198

The paper is devoted creation and research of algorithms of a speech dialogue control system.