

АВТОМАТИЗИРОВАННАЯ КОМПЬЮТЕРНАЯ СИСТЕМА ГОЛОСОВОГО УПРАВЛЕНИЯ АВТОМОБИЛЕМ

**А.П. Кравченко, профессор, д.т.н., Н.М. Крамарь, доцент, к.т.н.,
И.В. Морозов, ассистент, ВНУ имени В. Даля**

***Аннотация.** Рассмотрены отечественные автомобили и зарубежные аналоги. Отмечены недостатки, связанные с отсутствием вспомогательной электронной системы, служащей для повышения безопасности и удобства управления машиной. Предложена инновационная разработка комплексной системы голосового управления, обеспечивающая надёжность, удобство и простоту передвижения в автомобиле.*

***Ключевые слова:** голосовое управление, голос, речь, автомобиль, электроника, компьютерная система, речевой словарь.*

АВТОМАТИЗОВАНА КОМП'ЮТЕРНА СИСТЕМА ГОЛОСОВОГО КЕРУВАННЯ АВТОМОБІЛЕМ

**О.П. Кравченко, профессор, д.т.н., М.М. Крамарь, доцент, к.т.н.,
І.В. Морозов, асистент, ВНУ імені В. Даля**

***Анотація.** Розглянуто вітчизняні автомобілі і зарубіжні аналоги. Відмічено недоліки, пов'язані з відсутністю допоміжної електронної системи, що служить для підвищення безпеки і зручності управління машиною. Запропоновано інноваційну розробку комплексної системи голосового керування, що забезпечує надійність, зручність і простоту пересування в автомобілі.*

***Ключові слова:** голосове управління, голос, мова, автомобіль, електроніка, комп'ютерна система, мовний словник.*

AUTOMATED COMPUTER SYSTEM OF VEHICLE VOICE CONTROL

**A. Kravchenko, professor, dr. eng. sc., N. Kramar, associate professor, cand. eng. sc.,
I. Morozov, assistant, VNU after Dal'**

***Abstract.** Domestic cars and foreign analogues are considered. Failings are marked related to absence of the auxiliary electronic system which serves for the increase of safety and comfort of vehicle management. Innovative development of the complex system of vocal management which provides reliability, comfort and simplicity of movement in a vehicle is offered.*

***Key words:** voice control, voice, speech, vehicle, electronics, computer system, speech dictionary.*

Введение

С каждым днём появляются возможности усовершенствовать мир новыми технологиями и системами, делая жизнь более комфортной и удобной. Рассматривая повседневное исполнение транспортных средств, нельзя не отметить, что это далеко не совре-

менная техника, сложная в управлении, требующая дополнительных усилий и внимания. Автомобили, в которых основные бортовые функции управляются вручную и при этом отсутствует «умная» электроника, хотели бы упростить управление до подачи голосовых команд [1, 2].

Анализ публикаций

Распознавание команд голосом является актуальным вопросом передовых технологий. Несмотря на то, что современная наука достигла вершин прогресса, всё же в повседневной жизни очень трудно встретить транспортное средство, оснащенное системой распознавания голоса и управления с его помощью бортовыми командами. Передовые авто/концерны мира, такие как Ford, BMW, Mercedes, стремятся повысить безопасность и комфорт водителя, поэтому дают возможность управлять бортовой электроникой при помощи голоса («включи радио, включи следующую станцию, позвонить, номер») [3, 4].

Цель и постановка задачи

Целью работы есть повышение степени безопасности и простоты управления автомобилем. Разработка системы, позволяющей не отводить взгляд от дороги и не занимать руки водителя лишними манипуляциями. Голос является идеальным решением для управления бортовыми функциями. Отсутствие различных кнопок и необходимости их нажимать существенно снижает риск опасности отвлечь водителя во время управления автомобилем [4].

Принцип построения систем голосового управления

Голосовое управление имеет большие перспективы на автомобильном транспорте, однако ни отечественные, ни зарубежные автомобили не могут ощутить «силу слова» [4].

В связи с этим разработана компьютерная система голосового управления и контроля за бортовыми функциями автомобиля. Данная система сможет быть интегрирована в обычный автомобиль и при этом обеспечивать как и удобство во время езды, так и безопасность работы автомобиля.

Сегодня существуют два основных вида технологий распознавания голоса. Один из них – это распознавание речи, зависящее от диктора, т.е. пользователь должен сначала научиться систему распознавать его голос, и только после этого система может функционировать [5]. Второй – это распознавание речи, не зависящее от диктора, т.е. система способна распознать любую речь, независимо от того, кто

говорит. Системы распознавания речи, зависящие от диктора (голосозависимые), предназначены для одного конкретного пользователя. Другие способы распознавания, голосонезависимые системы разрабатываются для любого пользователя конкретного типа (например, американский английский). Это самые сложные в разработке и самые дорогие системы, а точность распознавания у них ниже. Однако эти системы более гибки. Адаптивные системы приспособляются к характеристикам нового диктора. Уровень их сложности лежит где-то посередине между голосонезависимыми и голосозависимыми системами. Системы распознавания изолированных слов работают с дискретными словами – в этом случае требуется пауза между словами. Это самая простая форма распознавания, так как в этом случае легко определяется конец речевого сигнала, а произношение слова не затрагивает другие слова. Поскольку в этих системах количество слов постоянно, то их легче проектировать. Системы распознавания непрерывного речевого сигнала работают с речевым потоком, в котором слова сливаются, т.е. не разделены паузой. Непрерывную речь обрабатывать гораздо сложнее по целому ряду причин, во-первых, трудно определить начало и конец слова. Вторая проблема – это коартикуляция. На звучание каждой фонемы влияет звучание соседних фонем, а на начало и конец слов влияют предыдущие и последующие слова. Распознавание непрерывной речи зависит также от скорости речи: с быстрой речью работать сложнее, чем с медленной [5, 6].

Размер словаря системы распознавания голоса влияет на степень сложности, требования к процедурам обработки и точность системы. Одним системам для работы необходимо всего несколько слов (например, только числа), а другие работают с очень большими словарями (например, диктофонные машины). Четких градаций объемов словарей нет. Но обычно словари классифицируются на:

- маленькие словари – это десятки слов;
- средние – сотни слов;
- большие словари – тысячи слов;
- очень большие словари – это десятки тысяч слов.

Есть еще две проблемы, тесно связанные с объемом словарей. Одна из них касается разработки и обеспечения доступа к специальным базам данных (словарям): необходимо

обеспечить ведение такой базы и возможность обновления данных для специальных групп пользователей различных профессий, например в медицинской или правовой сфере. Вторая проблема – это проверка грамматики. Ведь программы разрабатываются не только для записи продиктованных слов, но и включают функции проверки и исправления структур предложений.

Из вышесказанного следует, что системы распознавания голоса требуют огромных ресурсов, включая вычислительную мощность, память и сетевые возможности. С распознаванием голоса тесно связаны две технологии:

- цифровая обработка сигнала;
- распознавание образов.

Методы цифровой обработки сигнала обычно осуществляют преобразование, очистку и трансформацию акустического сигнала в цифровой формат данных и другие представления, которые могут непосредственно обрабатываться системой распознавания речи. Эти задачи включают фильтрацию шумовых сигналов, которые примешиваются к звуку при передаче акустических сигналов от воспринимающих устройств (микрофонов) или по сети. Методы распознавания образов используются при выделении и распознавании отдельных слов или предложений речевого потока или в некоторых случаях для идентификации говорящего. Кроме этого, понадобится лингвистическая теория – в ней заложены фундаментальные концепции и принципы распознавания речи и понимания языка [6].

Процесс распознавания голоса проходит в несколько этапов. На каждом из этапов для обработки речевого сигнала используется целый ряд различных методов. Процесс распознавания голоса можно разбить на три этапа:

- получение голосового сигнала и предварительная обработка речи;
- распознавание фонем и слов;
- понимание речи.

«Понять» речь – это самое трудное [7]. На этом этапе последовательности слов (предложения) должны быть преобразованы в представления о том, что хотел сказать говоривший. Хорошо известно, что понимание речи опирается на огромный объем лингвистических и культурных знаний. Большая часть систем распознавания голоса учитывает при этом знания о естественном языке и конкретные обстоятельства. Задача, связан-

ная с распознаванием голоса, – распознавание говорящего, т.е. процесс автоматического определения, «кто говорит» на основе входящей в речевой сигнал индивидуальной информации. При этом речь может идти об идентификации или о верификации говорящего. Идентификация – это нахождение в известном множестве контрольных фраз экземпляра, соответствующего манере данного диктора говорить. Верификация диктора – это определение идентичности говорящего: тот ли это человек? Технология распознавания диктора позволяет использовать голос для обеспечения контроля доступа; например, телефонный доступ к банковским услугам, к базам данных, к системам электронной коммерции или голосовой почте, а также доступ к секретному оборудованию. Обе технологии требуют, чтобы пользователь был «занесен в систему», т.е. он должен оставить образец речи, по которому система может построить шаблон. В процессе разработки предпринимались попытки создать и аппаратную реализацию систем распознавания голоса, но такие системы не показали высоких результатов. Чип поддерживает голосозависимое распознавание на базе словаря, хранимого в постоянном запоминающем устройстве чипа (ROM, read only memory). Словари голосозависимых систем хранятся вне чипа и могут быть загружены во время работы системы [8].

Для распознавания бортовых команд в автомобиле достаточно сложным является подавление окружающего шума работы двигателя, внешних звуков дороги и разговора пассажиров. Для этого в работе было предложено снимать звук не с общего микрофона, а с микрофонной гарнитуры, которая обеспечивает наиболее близкий контакт к голосовым связкам человека.

На рис. 1 приведена фотография установки компьютерной системы в автомобиль. Компьютер (ноутбук) в торпеде автомобиля снимает и подаёт сигналы бортовых функций всей машины, обрабатывает их и при помощи специально разработанного программного обеспечения координирует и запускает вспомогательные устройства.

Система, кроме управления основными командами, должна работать и как «умная» электроника. Гибкость программного обеспечения позволяет настроить работу всех



Рис. 1. Установленный компьютер в модернизированной торпеде автомобиля ВАЗ 2101

функциональных приборов логически. К примеру, самая простая защита от запуска стартера во время того, как двигатель уже работает. При прокрутке стартером включенного мотора изнашиваются шестерни, что в дальнейшем может нарушить запуск машины автомобиля. Данная система будет блокировать ошибочное действие человека и не запустит стартер, чем продлит ему срок службы. Естественно, что компьютерная система не смогла бы работать и анализировать происходящие процессы без дополнительных встроенных датчиков, которые являются «нервными окончаниями» всей системы.

Для водителя, который впервые сел за руль автомобиля, будет включен обучающий курс работы с данной системой, в ходе которого ему представится обучить компьютерную программу на распознавание его личного голоса.

Использование компьютера на борту автомобиля позволяет расширить функциональность системы дополнительными опциями. Замена зеркала заднего вида на инфракрасную камеру ночного видения, которая передает изображение на монитор. В систему заложено голосовое управление аудиосистемой (поиск музыки, регулирование громкости и т.п.). А также замена классической системы зажигания на микропроцессорную, роль блока управления которой также возложена на компьютер. Такая система обеспечивает стабильную работу двигателя и снижает расход топлива.

Вывод

Продолжается модернизация и исправление неточности работы системы, не решена окончательно и проблема отделения речевого

сигнала от шумового фона. В настоящее время пользователи систем распознавания голоса вынуждены либо работать в условиях минимального шумового фона, либо использовать микрофонную гарнитуру. Что касается того чтобы команда, случайно высказанная в слух, не запустилась, была добавлена активация по «имени» машины. Например, «Автомобиль, включи фары». Для этого в предыдущем варианте обычно надо нажать кнопку принятия команды. Решение этих проблем началось, и уже получены многообещающие результаты. Одна из долгожданных разработок в области распознавания голоса – это человеко-машинные диалоговые системы. Система «умеет» работать с непрерывным речевым потоком и предпринимать ответные действия. Диалоговый интерфейс в системе позволяет человеку разговаривать с машиной, создавать и получать информацию, решать свои задачи.

Литература

1. Dr.-Ing. D. Anselm, Ismaning; Zwei Jahre elektronische Wegfahrsperrung – Ergebnisse im Jahre 1996; *Elektronik im Kraftfahrzeug*, Tagung Baden-Baden, VDI-Berichte 1287, 1996. – Pp. 110–123.
2. Автомобильная энциклопедия 2008. – Издательство: Третий Рим Издательский дом (ООО), Издат. Дом «Третий Рим», 2007. – С. 7–50.
3. Joe Schurmann, Herbert Meier; TIRIS – Leader in Radio Frequency Identification Technology, *Texas Instruments Technical Journal* Vol. 10, No. 6.
4. W. Steinhagen, U. Kaiser; A Low Power Read / Write Transponder IC for High Performance Identification Systems, *Proceedings of ESSCIR' 94*, Ulm, Germany, September 20–22, 1994. – P. 256–259.
5. <http://www.sciencemag.org/>
6. <http://tehnology.com/>
7. <http://alife-oft.narod.ru>
8. <http://www.agentura.ru>

Рецензент: В.И. Клименко, профессор, к.т.н., ХНАДУ.

Статья поступила в редакцию 27 августа 2009 г.