

УДК 004.522

ПОТРЕБНОСТЬ В РЕЧЕВЫХ ФУНКЦИЯХ ИНТЕЛЛЕКТУАЛЬНЫХ САПР

Абрамов Г.А., Боргест Н.М., Коровин М.Д.

*Самарский государственный аэрокосмический университет имени академика С.П. Королева
(национальный исследовательский университет), г. Самара, Россия*

borgest@yandex.ru

maks.korovin@gmail.com

belg.ru@mail.ru

В работе рассмотрены основные принципы построения голосовых интерфейсов автоматизированных систем, вопросы распознавания и синтеза речи. С целью увеличения производительности труда и снижения временных затрат на обучение сложных программных приложений обосновывается необходимость применения голосовых технологий на примере разрабатываемого интеллектуального помощника проектанта в предметной области «самолетостроение».

Ключевые слова: голосовой интерфейс; распознавание; синтез; помощник проектанта; голосовые подсказки; голосовое управление.

Введение

Современные темпы развития пользовательских приложений приводят к усложнению программных интерфейсов. Наиболее ярко эта тенденция проявляется в «тяжелых» САД-системах, на освоение которых требуются значительные временные затраты. Одним из подходов решения указанной проблемы является применение справочных систем, построенных на основе речевых технологий. Внедрение в пользовательские инженерные инструменты «голосовых помощников» способствует наилучшему восприятию передаваемой информации и уменьшает время, необходимое на подготовку пользователя и на принятие решений. Реализация голосовых технологий в рабочем инструменте способствует увеличению производительности системы человек-машина.

1. Голосовой интерфейс

Пользовательский интерфейс представляет собой совокупность средств и методов, при помощи которых пользователь взаимодействует с различными машинами, устройствами и аппаратурой [Фисун, 2004].

Голосовой интерфейс – это разновидность пользовательского интерфейса, в которой обмен информацией проходит с помощью речи [Фролов, 2003]. Такой вид взаимодействия очень

привлекателен в различных областях, для которых характерна высокая интенсивность обмена данными между пользователем и информационной системой (ИС). Наиболее полно воспринимается та информация, которая передается одновременно несколькими методами, например визуальным и звуковым. Такой подход помогает обеспечить малые сроки освоения пользования системой и повысить эргономичность интерфейсов ИС в широком спектре предметных областей.

2. Краткая история развития речевых технологий

Работы по созданию систем автоматического распознавания речи ведутся с конца 50-х годов.

В 60-х годах появляются технологии, которые позволяют использовать имеющиеся голосовые метки для выполнения конкретных задач.

В 80-е годы компании Bell Laboratories и IBM приступили к исследованиям в направлении N-грамм (непрерывных последовательностей из N элементов заданного текста или речи), и спикерозависимых (зависимых от свойств звукового входного сигнала) тренируемых систем.

В 90-е годы начинает развиваться технология нейронных сетей. Данный подход увеличивает степень точности распознаваемой речи и ведет к ускорению обработки информации [Petersche, 1998].

В 2006 году Microsoft выпускает операционную систему с полноценной поддержкой функции распознавания речи Windows Vista.

В 2011 году была разработана программа Siri – самообучаемый персональный помощник с голосовым интерфейсом.

Активное внедрение элементов голосового управления в популярные пользовательские приложения позволяет надеяться на скорую готовность речевых технологий к применению в более сложных предметных областях, например, в тяжелых CAD.

3. Распознавание речи

Для успешной реализации голосового интерфейса в определенной предметной области, необходимо принимать во внимание наличие специфической лексики. Уровень развития вычислительной техники и речевых технологий в настоящий момент не позволяют пока реализовать диалог на естественном языке между ИС и пользователем. Это обусловлено такими факторами, как:

- ограниченность словаря ИС;
- особенности дикции оператора и различные скорости произношения;
- недостаточное качество работы аппаратных интерфейсов ввода речевой информации;
- наличие посторонних шумов, влияющих на точность получения команд;
- отсутствие на сегодняшний день технологий семантического анализа текста, готовых к применению в пользовательских приложениях.

Для устранения ограничений, связанных с возможностью некорректной выдачи и восприятия информации, были разработаны специальные языки, например, упрощенные технические языки, главной целью которых является обеспечение однозначности передаваемой информации [Степура, 2013]. Примером такого языка является Simplified Technical English (STE) - упрощенный технический английский язык (<http://www.simplifiedenglish.net/>) или УТР – упрощенный технический русский (<http://www.technorus.ru/>). Переход к упрощенному языку обеспечивает снижение погрешности машинного восприятия текста на 20-25%. STE был разработан при участии компании Boeing для сокращения номенклатуры и содержимого выпускаемой документации, обеспечения полного и однозначного понимания информации.

При использовании упрощенных языков, положительный эффект достигается за счет:

- использования слова только в одном определенном смысле;
- использования простых предложений;
- запрещения в простом предложении постановки подлежащего в конец предложения;
- запрещения изменения инвертированного порядка частей составного сказуемого;

- запрещения постановки прямого дополнения перед сказуемым;

- запрещения использования причастных оборотов в качестве синтаксических оборотов.

Использование упрощенного технического языка при работе с голосовым интерфейсом ИС позволяет повысить качество работы за счет уменьшения количества ошибок распознавания.

На сегодняшний день существуют два принципиально отличающихся метода распознавания звукового потока. Это метод распознавания голосовых меток и метод распознавания лексических элементов.

Первый подход предполагает распознавание фрагментов речи по заранее записанному образцу. Этот подход широко используется в системах, предназначенных для исполнения заранее записанных речевых команд.

Второй подход реализуется выделением из потока речи лексических элементов — фонем и аллофонов (реализация фонемы, её вариант, обусловленный конкретным фонетическим окружением), которые затем объединяются в слоги и морфемы. Применение этого подхода в интерактивных системах осложняется его высокой ресурсоемкостью.

Вне зависимости от реализуемого подхода, распознавание состава речи происходит в три этапа:

- выделение лексических элементов речи;
- выделение слогов и морфем;
- выделение слов, предложений и сообщений.

После завершения распознавания речи, система обрабатывает полученные команды и, в случае необходимости, синтезирует голосовые комментарии.

4. Синтез речи

Системы синтеза речи подразделяется на 4 основных типа [Феррейра, 2012] :

- параметрический синтез – синтез, где речевой сигнал представляется набором небольшого числа непрерывно изменяющихся параметров. При данном подходе, выходной сигнал, в зависимости от количества параметров, может быть очень высокого качества. Недостатком является невозможность применения параметрического синтеза в случаях произвольных, не подготовленных диалогов;

- конкатенативный или компиляционный (компилятивный) синтез – синтез, который сводится к составлению сообщения из предварительно записанного словаря исходных элементов. Размер элементов синтеза не меньше слова;

- синтез по правилам – синтез, реализуемый путём моделирования речевого тракта человека. В процессе синтеза речи фонемы выводятся последовательно через определённый временной интервал;

- предметно-ориентированный синтез –

синтез, который компилирует слова, записанные заранее, а также фразы для создания полных речевых сообщений. Этот тип синтеза используется в приложениях, где многообразие текстов системы будет ограничено определенной темой.

5. Применение речевых технологий в автоматизированных САПР

В силу того, что разрабатываемый помощник проектанта реализуется в узкоспециализированной предметной области, в частности, предварительное проектирование самолета, функциональный лексикон пользователя ограничен тезаурусом предметной области.

Тезаурус является одним из действенных инструментов описания отдельных предметных областей. Он состоит из контролируемого, но при этом изменяемого словаря терминов, между которыми устанавливаются смысловые связи. Возможность формализации семантики речевых команд через взаимосвязь лексических единиц с множеством структурированных понятий предметной области является отличительной особенностью тезаурусов от толковых словарей. В тезаурусе описание понятий предметной области осуществляется систематизировано и направлено на выявление отношений иерархического и неиерархического типов. Таким образом, можно сказать, что тезаурус обеспечивает удовлетворение одного из важных требований при использовании САПР в практике проектирования – наличие полной информации о проектируемом объекте. Тезаурус является важным информационным ресурсом при обработке и анализе данных речевой системой [Боргест, 2012].

В настоящий момент речевая подсистема в интеллектуальном помощнике проектанта исполняет роль интерактивного тьютора. Задачей речевой компоненты является озвучивание для

пользователя алгоритма работы с графическим интерфейсом помощника, а так же принятых в автоматическом режиме ключевых решений на этапе предварительного расчета проекта самолета. Это позволяет упростить контроль над системой и повысить удобство работы пользователя.

Перспективным направлением развития речевого интерфейса интеллектуального помощника проектанта видится реализация голосового управления на основе тезауруса предметной области и статистической базы данных голосовых команд, что позволило бы освободить пользователя от необходимости ручного ввода параметров задачи.

Исходя из этого, можно утверждать, что для автоматизированной системы поддержки принятия конструкторских решений целесообразно применить предметно-ориентированный синтез.

Учитывая сложность интерфейса автоматизированных приложений ведущие разработчики САПР ставят перед собой задачу разработки системы, способной общаться с пользователем на тривиальном языке, что позволило бы свести к минимуму время, необходимое для освоения интерфейса программы. Исследователи из Университета Гонконга предложили схему реализации общения между САД системой и проектантом, представленную на рисунке 1 [Кou, 2010]. Авторы этого исследования предлагают создать базу знаний, специфических для предметной области слов и выражений, характерных для конечных пользователей. Такой подход позволит повысить точность распознавания, так как система анализирует характерные фразы, которым назначены конкретные действия. Успехи, полученные при применении системы, демонстрируют перспективность такого подхода в специализированных системах, предназначенных для работы в областях, где традиционно часто применяются специальные термины и жаргонизмы.

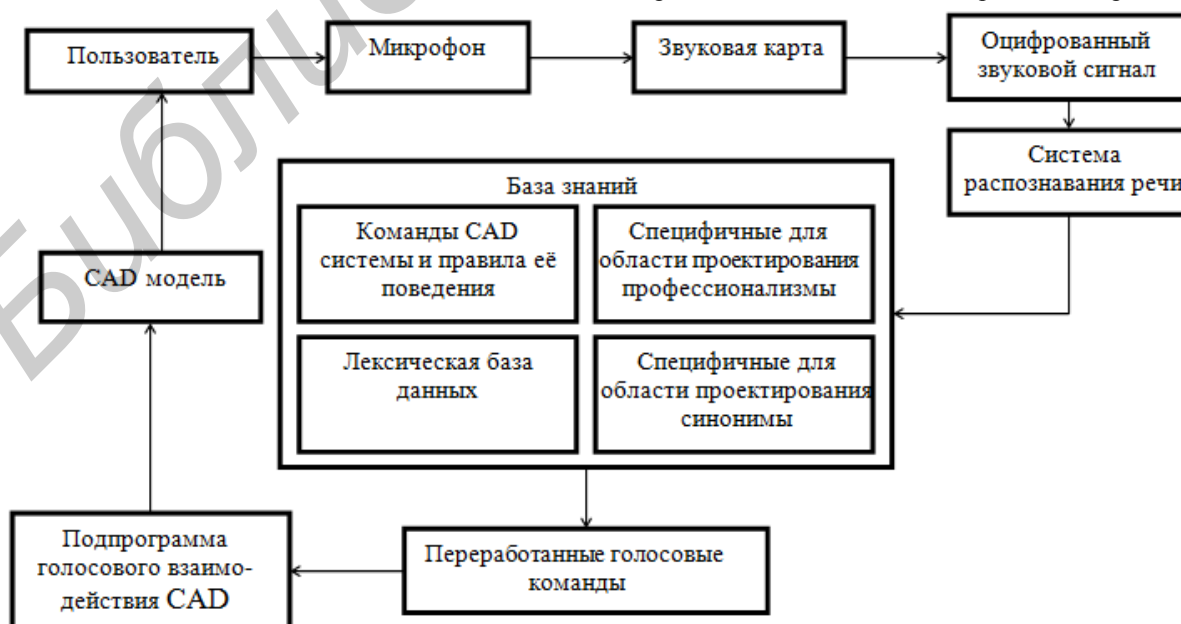


Рисунок 1 - Модель голосового управления САД системой на тривиальном языке

Заключение

Усложнение графических интерфейсов современных CAD продуктов вследствие расширения их функциональных возможностей диктует необходимость выработки новых подходов к реализации взаимодействия между системой и пользователем. Одним из наиболее перспективных способов снизить время, потребное для освоения ИС, является внедрение речевых технологий, способных к интерактивной выдаче справочной информации, а в перспективе – к реализации диалога с пользователем на естественном языке.

Библиографический список

[Боргест и др., 2012] Боргест Н.М., Чернов Р.В., Шустова Д.В. Разработка интерфейса интеллектуального помощника проектанта / Н.М. Боргест, Р.В. Чернов, Д.В. Шустова // материалы международной научно-технической конференции OSTIS-2012. БГУИР. 2012

[Соловов, 2006] Соловов А.В. Электронное обучение: проблематика, дидактика, технология / А.В. Соловов // Самара: «Новая техника», 2006. – С. 464

[Степура, 2013] Степура Л. В. Автоматическое реферирование текстовой информации на основе моделирования ситуативных связей между понятиями предметной области / Л.В. Степура // Информатика, 2013, № 39

[Сулейманов, 2011] Сулейманов Д.Ш. Двухуровневый лингвистический процессор ответных текстов на естественном языке//Материалы международной научно-технической конференции OSTIS-2011, Минск, БГУИР, 2011 [Феррейра, 2012] Феррейра Опасо Е.В. Интерактивный голосовой интерфейс / Е.В. Феррейра Опасо // <http://sibac.info/index.php/2009-07-01-10-21-16/3955-2012-10-02-04-01-00> / актуально на 02.10.2012

[Фисун и др., 2004] Фисун А.П., Гращенко П.А. Теоретические и практические основы человеко-компьютерного взаимодействия: базовые понятия человеко-компьютерных систем в информатике и информационной безопасности / А.П. Фисун, П.А. Гращенко // Орел, ОГУ, 2004. – С. 169.

[Фролов и др., 2003] Фролов А.В., Фролов Г.В. Синтез и распознавание речи. Современные решения / А.В. Фролов, Г.В. Фролов // <http://www.frolov-lib.ru/books/hi/index.html/> актуально на 15.10.2013

[Kou, 2010] Kou X.Y., Xue S.K.. Knowledge-guided inference for voice-enabled CAD / X.Y. Kou, S.K.. Xue // *Tan Computer-Aided Design Volume*. – 2010 - P. 545-557.

[Markowitz, 2000] Markowitz A.J. Using speech recognition / A.J. Markowitz // Prentice Hall PTR. - 2000.

[Petsche, 1998] T. Petsche, A. Marcantonio, C. Darken, S. J. Hanson, G. M. Huhn, and I. Santoso, "An autoassociator for on-line motor monitoring," in *Industrial Applications of Neural Networks*, F. F. Soulie and P. Gallinari, Eds, Singapore: World Scientific, 1998, pp. 91–97.

THE NEED FOR VOICE CONTROL IN INTELLECTUAL CAD

Abramov G.A., Borgest N.M., Korovin M.D.

*Samara State aerospace University
(national research university),
Samara, Russian Federation*

borgest@yandex.ru

maks.korovin@gmail.com

belg.ru@mail.ru

The paper discusses the basic principles of voice interfaces implementation into high-tech automated systems. The principles of speech recognition and synthesis are described. The use of voice technology in complex user applications is given. Increase in productivity and time spent on learning to use the software reduction is described based on the example of the developed intelligent designer's assistant in the preliminary design of an aircraft domain.

Introduction

The current pace of development of custom applications dictates the increase of their UIs complexity. One of the solutions to this problem is to use a voice technology based assistant. Introduction of custom engineering tool "voice assistant" promotes the best perception of the transmitted information and reduces the time needed to educate the user.

Main Part

The implementation of voice technology into the intellectual designers assistant is described:

- Advantages and disadvantages of voice technology in CAD software is given.
- Trend for the increase of voice systems usage in software interfaces is observed.
- Speech recognition systems are discussed from the point of their applicability to intellectual designers assistant.
- Speech synthesis systems are examined for their applicability to intellectual designers assistant.

Conclusion

The increasing complexity of modern CAD software due to the advance in their functionality calls for a new approach to UI. One of the most promising technologies that can drastically decrease the amount of time, needed to educate the user is the speech-aided UI element implementation. It can serve as an interactive reference system and potentially might be able to support a machine-user dialog using trivial language.