



Литература

1. Галиев Ш.И., Карпова М.А. Оптимизация многократного покрытия ограниченного множества кругами // Журнал вычислительной математики и математической физики. 2010. Т. 50. № 4. С. 757–769.
2. Лебедев П.Д., Бухаров Д.С. Аппроксимация многоугольников наилучшими наборами кругов // Известия Иркутского государственного университета. Серия: Математика. 2013. № 3. С. 72–87.
3. Лебедев П.Д., Успенский А.А., Ушаков В.Н. Алгоритмы наилучшей аппроксимации плоских множеств объединениями кругов // Вестник Удмуртского университета. Математика. Механика. Компьютерные науки. 2013. Вып. 4. С. 88–99.
4. Heppes A. Covering a planar domain with sets of small diameter // Periodica Mathematica Hungarica. 2006. Vol. 53. P. 157–168.
5. Забелин С.Л., Фроловский В.Д. Исследование метаэвристических алгоритмов для решения задач оптимального геометрического покрытия // Перспективные информационные технологии в научных исследованиях, проектировании и обучении. 2012. С. 127 – 136.
6. Забелин С.Л., Фроловский В.Д., Жеголко К. В. Разработка и исследование генетического алгоритма для автоматизации проектных процедур оптимизации геометрического покрытия // Вестник ТГТУ. 2015. Т. 21. №2. С. 257 – 265.
7. Кочкаров А.А., Яцкин Д.В. Алгоритм поиска оптимального расположения сенсоров для решения задачи мониторинга пространства // Программные продукты и системы. 2016. Т. 29. №3. С. 60 – 66.
8. Кутузов О. И., Татарникова Т. М. Решение одной задачи размещения сенсорных устройств в сетях интернета вещей // Известия СПбГЭТУ «ЛЭТИ» № 6/2018. С 15 – 20.

А. Н. Спиркин

СУБВОКАЛЬНОЕ УПРАВЛЕНИЕ РОБОТИЗИРОВАННЫМ МЕХАНИЗМОМ

(Пензенский государственный университет)

В настоящее время автоматическое распознавание речи является динамично развивающимся направлением в области создания интерфейсов взаимодействия с техникой. За последние полвека в данной области достигнуты значительные успехи – имеется множество коммерческих приложений, которые делают вложения в данную область оправданными и выгодными. Следует отметить, по крайней мере, три основных преимущества использования речевого интерфейса:



1) для общения с компьютером человеку нужно пройти специальную подготовку. В то же время речь – это естественный интерфейс для любого, даже неподготовленного человека.

2) речь сама по себе никак механически не привязана к компьютеру и может быть связана с ним через системы коммуникаций, например, телефон, что дополнительно расширяет круг потенциальных пользователей компьютеров и делает речевой интерфейс идеальным средством для создания систем массового информационного обслуживания;

3) речевой интерфейс позволяет обращаться к компьютеру в полной темноте, с закрытыми глазами, в условиях занятости рук рычагами управления, с завязанными руками и в другой экстремальной обстановке. Это свойство даёт оперативность и мобильность общения, освобождение рук и разгрузку зрительного канала восприятия при получении информации. Это исключительно важно, например, для диспетчера большой энергетической системы или пилота самолёта и водителя автомобиля [1].

Как и любой человеко-машинный интерфейс, речевой интерфейс должен удовлетворять таким требованиям как естественность, интуитивность, дружелюбность, согласованность и адаптивность. Однако голосовое управление имеет определенную специфику, что предъявляет к речевому интерфейсу дополнительные требования [2]. Основные проблемы, которые стоят перед разработчиками систем речевых интерфейсов, заключаются в следующем:

- необходимость использования больших вычислительных ресурсов – речевой интерфейс, распознающий фразу слишком долго, может не только не повышать эффективность работы оператора, но и понижать производительность работа в целом;

- подавление внешнего шума – речевые интерфейсы очень чувствительны к шуму окружающей среды, любые внешние помехи препятствуют правильному распознаванию речи;

- проблемы распознавания слитной речи, так как система распознавания речи – это сложный процесс согласования данных, любая, даже незначительная ошибка может привести к получению неверного результата;

- чувствительность речевых интерфейсов к речевым оборотам самого диктора.

Речевой интерфейс может работать в двух режимах: дикторозависимом и дикторонезависимом. В дикторозависимом режиме производится настройка системы на определенного диктора, и предполагается, что будут распознаваться слова, произнесенные только этим диктором. В дикторонезависимом режиме такая настройка не требуется, и система должна в идеале «понимать» любого диктора [2]. Последний режим, очевидно, удобнее в использовании, но точность распознавания в нем всегда меньше, чем в первом.

Голосовое управление происходит по следующим этапам — сегментация, распознавание речи, отклик на требуемую задачу. Обработка речи начинается с определения уровня помех и искажений, полученного речевого сигнала. Далее в сигнале выделяются участки, которые содержат речь, и происходит оценка



информации о формах слов. Этот этап называется сегментацией. Далее параметры речи поступают в декодер, где сопоставляются входные речевые потоки информации с потоками, хранящимися в акустических и языковых моделях устройства. Там же определяется некоторая наиболее вероятная последовательность слов, которая в итоге будет являться результатом. Это этап распознавания речи. Последнее — отклик на требуемую задачу. На этом этапе процессор, встроенный в используемое устройство, обрабатывает полученную команду. И если команда совпадает с той, что хранится в его встроенной базе данных, то устройство выполняет требуемую задачу, если же команда отсутствует, то используемое оборудование оповестит о невозможности ее выполнения.

Особенно остро проблема распознавания речи стоит необходимости управлять роботизированными механизмами в условиях воздействия внешней среды, например, космосе, под водой, во время проведения боевых операций. Чтобы решить указанные проблемы речевого интерфейса, автором предлагается использовать субвокальное распознавание команды.

Субвокальное распознавание – процесс регистрации субвокализации посредством электромиографических (ЭМГ) электродов движения гортани и других мышц, участвующих в артикуляции речи, и преобразования результатов в цифровой выход, звуковой или текстовый. Данный метод похож на распознавание речи, за исключением того, что субвокализация регистрируется в процессе молчания.

Обобщенная структура речевого интерфейса приведена на рисунке 1. Речевого интерфейс представляет собой аппаратно-программный комплекс, в котором можно выделить три основных уровня:

- задающий уровень управления, включающий аппаратные средства получения речевого сигнала;
- логический уровень управления, содержащий алгоритмы распознавания речевой команды и формирования управляющих команд для исполнительного уровня;
- исполнительный уровень, включающий устройства управления отдельными звеньями промышленного робота.

Особый интерес имеет второй уровень управления – логический, который отвечает за распознавание и идентификацию речевой команды.

В настоящее время лучшим программным средством распознавания речи стали искусственные нейронные сети, что-то по типу BiLSTM [3], на которых построены все современные сервисы распознавания голоса. Искусственная нейронная сеть позволяет с высокой точностью понимать слова, а также предсказывать наиболее вероятное слово в рамках контекста, если оно не было распознано. Нейронная сеть выделяет в зарегистрированном сигнале отдельные фонемы (минимальные смыслоразличительные единицы языка). Фонема не имеет самостоятельного лексического или грамматического значения, но служит для различения и отождествления значимых единиц языка (морфем и слов). Далее, после многократного анализа, нейронная сеть четко выделяет определенные фонемы, а их текстовая запись сравнивается с базой слов нейросети и



затем превращается в распознанное слово [4]. После чего результат работы нейронной сети можно перевести в управляющую команду для исполнительного уровня, который представляет собой роботизированный механизм.

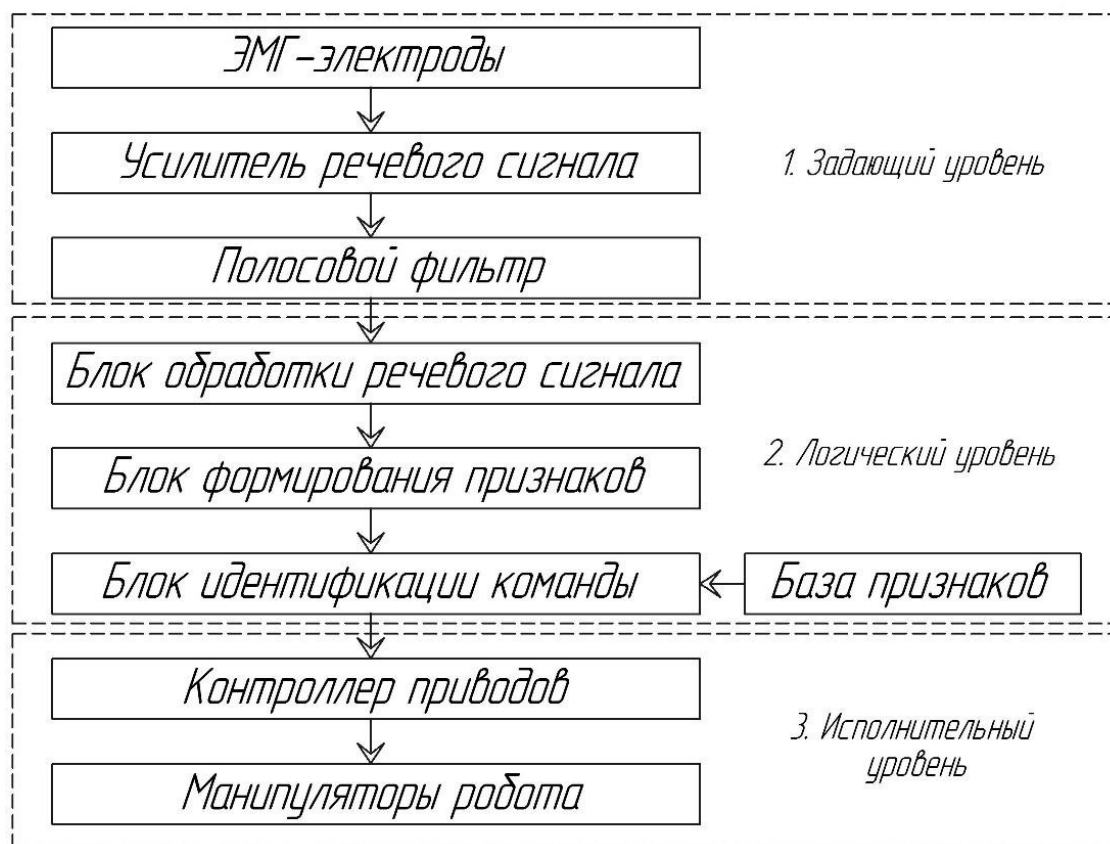


Рисунок 1 – Обобщенная структура речевого интерфейса

Таким образом, использование субвокального управления роботизированными механизмами целесообразно применять там, где очень сложно или невозможно услышать речь, например, для астронавтов, подводного спецназа, лётчиков-истребителей и аварийных работ в шумной окружающей среде.

Литература

1. Потапова Р.К. Речевое управление роботом: лингвистика и современные автоматизированные системы. – М.: Радио и связь, 1989. – 248 с.
2. Бройнль Томас Встраиваемые робототехнические системы. Проектирование и применение мобильных роботов со встроенными системами управления / Томас Бройнль. – М.: Институт компьютерных исследований, 2012. – 520 с.
3. S. Hochreiter, J. Schmidhuber, Long short-term memory. Neural Comput. 1997, vol. 9, pp. 1735–1780.
4. Тампель И.Б., Карпов А.А. Автоматическое распознавание речи. Учебное пособие. – СПб: Университет ИТМО, 2016. – 138 с.