

УДК 004.934.2+004.52

К ВОПРОСУ ОБ ИСПОЛЬЗОВАНИИ ЭМОЦИОНАЛЬНОЙ ОКРАШЕННОСТИ КОМАНДЫ ПРИ ГОЛОСОВОМ УПРАВЛЕНИИ РОБОТОМ ¹

И.П. Карпова (*karpova_ip@mail.ru*)

Национальный исследовательский университет
«Высшая школа экономики», НИЦ «Курчатовский
институт», Москва

М.А. Ровбо (*rovboma@gmail.com*)

НИЦ «Курчатовский институт», Москва

В докладе рассматривается управление роботом с помощью голосовых команд, которые преобразуются в текст и представляются в виде набора слов. Предлагаемый метод распознавания команд основан на сопоставлении текста поступившей команды со словарем и поиске слов-модификаторов, отражающих эмоционально-экспрессивную окрашенность команды и влияющих на приоритет ее выполнения.

Ключевые слова: мобильный робот, голосовое управление, эмоционально-экспрессивная окрашенность команды

Введение

Человеко-машинные интерфейсы управления роботом, основанные на голосовой связи, являются естественным для человека способом отдавать роботу различные команды. В данном докладе рассматривается работа с голосовой командой, преобразованной в текст и представленной в виде набора слов. Основная задача – точное и быстрое распознавание команды, отданной пользователем. Важными элементами голосового управления являются психологические аспекты [Ющенко, 2009], а также естественность передачи вместе с командой дополнительной информации

¹ Работа выполнена при финансовой поддержке РНФ (проект №16-11-00018) и РФФИ (проект офи_м №17-29-07083).

(такой, как приоритет). Приоритет можно связать с эмоционально-экспрессивной окрашенностью команды, отдаваемой оператором: чем она более явно выражена, тем больший приоритет имеет команда.

В настоящее время активно эксплуатируются роботы, которые обладают узкоспециализированным функционалом: транспортные роботы, роботы-пожарные и т.п. Управление такими системами может осуществляться с помощью ограниченного подмножества естественного языка, который робот должен распознавать, что обусловлено наличием конкретного перечня выполняемых действий. Например, в [Sivakumar et al., 2013] описана система голосового управления роботизированной инвалидной коляской, которая распознает всего 5 команд (вперед, назад, налево, направо, продолжать). Для голосового управления роботами с ограниченным набором команд не требуется большой словарный запас и можно использовать относительно простую систему распознавания, способную распознавать изолированные слова и короткие фразы [Ondáš et al., 2013]. Для этого нет необходимости проводить серьезный морфологический, синтаксический и семантический анализ, поэтому и в нашем случае эти этапы можно опустить.

При организации управления роботом широко распространен подход, основанный на эмоциональной архитектуре систем управления роботом. Эмоция рассматривается как механизм внутренней оценки ситуации роботом и как способ влияния на поведение робота. Например, в [Gadanh et al., 2003] описывается подход, при котором робот в процессе обучения с подкреплением использует эмоциональную оценку ситуации для перехода от одной поведенческой процедуры к другой.

Существуют роботы, которые умеют определять эмоциональный настрой взаимодействующего с ними человека и, в зависимости от этого, меняют свое поведение. Голосовой интерфейс таких систем, в основном, ориентирован на определение эмоций из акустических характеристик речи человека. Более того, в таких интерфейсах обычно выявляются сложные человеческие эмоции (гнев, радость, раздражение и т.д.), поэтому используются комплексные подходы: определение эмоций по голосу и выражению лица, по голосу и жестам и т.д. [Alonso-Martín et al., 2013].

Более близкой к нашей тематике является работа [Schuller et al., 2008]. В инвазивных операциях хирург ориентируется по изображению, получаемому с камеры, которой управляет робот с помощью голосовых команд хирурга. Для повышения адекватности управления в контур взаимодействия человек-робот включен учет эмоций. Голос хирурга сопоставляется с 3035 записями базы данных спонтанной эмоциональной речи в реальных хирургических операциях. Если система обнаруживает раздражение в голосе хирурга, активизируется диалоговое окно обратного вызова для подтверждения того, что камера направлена правильно.

Для определения эмоционального характера текста обычно используются специальные словари эмоционально окрашенных слов [Tausczik et al., 2010]. В таких словарях приводятся слова, отражающие степень проявления чувств и переживаний (экспрессию), например: уверенность (всегда, никогда), сомнение (предположительно, я думаю, возможно) и т.д. Для учета эмоционально-экспрессивной окраски команды также воспользуемся подобным словарем.

1. Постановка задачи

Пусть имеется мобильный робот, оснащенный некоторым количеством датчиков и системой восприятия голосовых команд. Рассматриваемый робот хранит задачи в очереди, каждой задаче соответствует процедура в виде конечного автомата. Когда заканчивает свою работу один автомат (т.е. задача выполнена или отменена), то продолжает работу следующий. Выполнение каждой процедуры инициируется внешней командой, которая представляет собой пару: имя выполняемой процедуры и приоритет команды. В каждый момент времени робот может выполнять только одну процедуру, поэтому поток команд оператора задает последовательность выполнения роботом тех или иных поведенческих процедур.

Местоположение новой команды в очереди определяется ее приоритетом. Это означает, что, в зависимости от приоритета, новая команда может быть поставлена в очередь как некоторое отложенное действие, может прервать выполнение текущей процедуры и активизировать новую, более приоритетную, а может и вовсе отменить выполнение текущей процедуры. Т.о., если приоритет пришедшей команды максимален, то выполнение текущей команды должно быть приостановлено, а управление передано новой команде, помещаемой в стек. Если приоритет новой команды меньше приоритета текущей, то команда помещается внутрь очереди в соответствующую позицию.

2. Выделение императивных форм и эмоционального компонента

Целью системы голосового управления является удобство и надёжность, в частности, важно минимизировать число ошибочных распознаваний, а также сопоставлять командам соответствующий приоритет исполнения. Это можно реализовать, учитывая эмоционально-экспрессивную окрашенность команд, отдаваемых роботу. Такой подход позволит естественным для человека образом отделять высокоприоритетные задачи от низкоприоритетных. Для этого предлагается использовать специальные ключевые слова в командной фразе (например, «немедленно», «сейчас же» или «потом»), которые

позволяют определить необходимость повышения или понижения приоритета команды по сравнению с её значением по умолчанию.

2.1. Эталонное множество команд и базовые приоритеты

Язык команд выбран весьма простой и ограниченный, задаваемый заранее определённым списком, соответствующим возможностям робота. Простота языка важна также для реализации алгоритмов на бортовой вычислительной системе с ограниченными ресурсами.

Множество доступных команд и соответствующих им фраз задаётся в виде словаря, включающего базовый приоритет команды (тот приоритет, который у неё есть до применения модификаторов). Фраза состоит из одного или нескольких слов, разделённых пробелами. Помимо основного словаря, имеется список слов-модификаторов, которые придают команде эмоциональную окрашенность. Эти модификаторы обрабатываются отдельно, перед тем как искать фразу в словаре команд. Их эффект (понижение или повышение приоритета) запоминается, после чего они удаляются из фразы, и она поступает на дальнейшую обработку. Значения приоритетов команды p_1 и модификатора p_2 (ключевого слова) должны лежать в интервале $[0;1]$. Итоговый эффективный приоритет p команды рассчитывается по общеизвестной формуле:

$$p = p_1 + p_2 - p_1 p_2. \quad (1.1)$$

Полученная в результате величина p монотонна и имеет ту же область значений $[0, 1]$.

Помимо эмоциональной окрашенности, управляющей приоритетом, также необходимо отдельно выделять случаи, когда требуется отменить выполнение предыдущих команд. Таким образом, базовый механизм сопоставления команд расширен маркерами приоритетов и отменяющими командами.

2.2. Метод сравнения множеств и списков. Для выбора команды из списка команд необходимо определить степень сходства с множествами слов, представляющих командную фразу в словаре, а также распознать модификаторы и обработать их приоритеты. Для решения этих задач применяется метод сравнения множеств, описанный в [Карпов и др., 2000; Карпова, 2001], и итерационная оценка приоритета. За область значений степени сходства δ принимается интервал $[0, 1]$: $\delta=1$ для идентичных множеств, $\delta=0$ для множеств, состоящих из разных элементов.

Степень сходства множеств вычисляется как величина, обратная к расстоянию r между множествами (r определено на интервале $[0, 1]$):

$$\delta = 1 - r. \quad (2.1)$$

Расстояние должно удовлетворять следующим условиям:

- Неотрицательность: $|AB| \geq 0, \forall A, B;$
- Идентичность: $|AB| = 0 \Leftrightarrow A=B;$
- Симметричность: $|AB| = |BA|, \forall A, B.$

Расстояние называется метрикой [Хачумов, 2012], если для него выполняется неравенство треугольника: $|AB| + |BC| \leq |AC|, \forall A, B, C.$

Пусть даны два множества A и B , состоящие из элементов базового множества R . Элементы множества могут быть скалярными величинами, множествами и списками. Расстояние между множествами скалярных величин может быть определено как:

$$r = 1 - \frac{|A \cap B|}{|A \cup B|}, \quad (2.2)$$

и тогда степень сходства таких множеств δ^S можно оценить по формуле:

$$\delta^S = \frac{|A \cap B|}{|A \cup B|}. \quad (2.3)$$

Рассмотрим случай, когда элементами множества являются множества. Для случая, когда $A=(a_1, a_2, a_3, \dots)$ – словарь команд, а $B=(b_1)$ – фраза, полученная от пользователя, определение соответствующей b_1 команды a_k и ее приоритета может быть выполнено в два этапа:

Этап 1. $b_1 \in B$ сравнивается со всеми элементами множества A (по правилам определения сходства множеств, (2.3)). На основе полученных значений вычисляется величина δ_i , которая рассматривается как степень уверенности в том, что i -й элемент множества A соответствует элементу b_1 . Индекс команды, имеющей максимальную степень сходства с распознанным текстом, определяется следующим образом:

$$k = \operatorname{argmax}(\delta_i)$$

Этап 2. Приоритет распознанной команды a_k с базовым приоритетом p_1 при применении модификатора с весом p_2 рассчитывается по формуле (1.1).

Если к команде применяется несколько модификаторов, то приоритет определяется итеративно с помощью формулы (1.1). Модификатор добавляется в двух случаях: во входной фразе b_1 распознано ключевое слово-модификатор, либо найден повтор текущей словарной команды в этой же фразе. Во втором случае для первой итерации $p_1=p_2$, затем к полученному значению итерационно добавляются веса модификаторов.

Таким образом, голосовая команда рассматривается как множество слов, которое сопоставляется с командами из словаря с помощью формулы (2.3). А приоритет команды определяется с учетом слов-модификаторов и повторов команд по формуле (1.1).

3. Эксперименты

Для валидации предложенного метода был проведён качественный тест с двумя конфигурациями системы: без использования приоритетов (новая команда помещается всегда в верхушку стека) и с использованием приоритетов (в зависимости от ключевых слов, команда переносится в соответствующее место стека). На реальном роботе проверялись возможность и удобство управления роботом на простой задаче, которая заключалась в обнаружении объекта на полигоне и его перемещения (требовалось подвинуть кубик, толкая его роботом). Для этого у робота имеется набор поведенческих процедур и соответствующих им команд: свободное блуждание и поиск объекта, взаимодействие с объектом, движение вперед, "сон" (режим ожидания на месте), отмена текущей команды, сброс всего стека команд. Обнаружение объекта производилось с помощью специального модуля видеообработки [Московский и др., 2018]. Тестовая платформа включала мобильного робота (Рис. 1) и отдельный компьютер, который управлял роботом по беспроводной связи. Архитектура системы обработки голосовой команды (Рис. 2) создана на основе платформы ROS [Quigley et al., 2009].

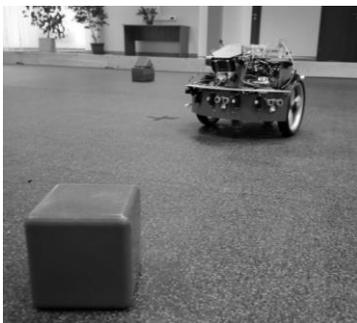


Рис. 1. Мобильный робот и целевой объект

Узел *speech_to_text* обеспечивает приём звуковых данных, разбиение их на фразы, исходя из пауз в речи, пересылку этих коротких звуковых записей Google или Yandex сервису распознавания голоса с помощью их API и отправку полученного текстового сообщения узлу *text_command_recognition*. Этот узел с помощью описанного выше метода сопоставляет текст с одной из команд в словаре и оценивает ее приоритет. После этого соответствующий конечный автомат добавляется в очередь команд и выполняется узлом *state_machine_processor* на основе данных об объектах, обнаруженных дальномерами и модулем компьютерного зрения.

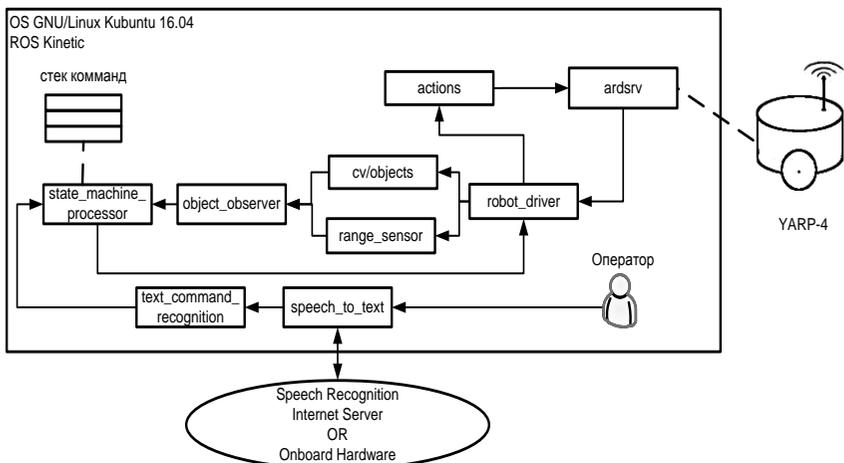


Рис. 2. Архитектура системы обработки голосовой команды

Список ключевых слов, влияющих на поведение, приведен в Табл. 1.

Табл. 1. Модификаторы команд

Тип	Приоритет	Слово
Обращение	Не меняет	"робот"
Приказ	Повышает	"немедленно"
Неотложное выполнение	Повышает	"сейчас"
Приказ	Повышает	"быстро"
Неуверенность	Понижает	"пожалуй"
Неуверенность	Понижает	"наверное"
Отложенное выполнение	Понижает	"потом"
Просьба	Понижает	"пожалуйста"

Полигон представлял собой область, ограниченную деревянными препятствиями достаточной высоты для обнаружения датчиками робота. На полигоне располагался набор различных разноцветных препятствий и целевой объект — зеленый кубик, который требовалось подвинуть. Работа оператора с вариантами системы без приоритетов и с ними выглядела следующим образом:

- Стек без приоритетов: [Сон, Вперед, Поиск] -> [Очистка стека] -> [Толкай]
- Стек с приоритетами: [Сон, Вперед, Поиск] -> [Сон, Вперед, Толкай, Поиск]

В ходе эксперимента робот выполнял все задачи последовательно или с учетом приоритетов, определяемых по словам-модификаторам.

Эксперимент показал работоспособность предложенного метода голосового управления.

Заключение

Была проверена идея о полезности определения приоритета на основе выявления эмоционально-экспрессивной окрашенности команды оператора. Была подтверждена работоспособность алгоритма итерационной оценки приоритета команд, распознанных с помощью метода сравнения множеств. Без учета приоритета оператору приходится, задав последовательность команд, при необходимости изменений отменять все команды и задавать их заново. В то время как с системой определения приоритетов оператор может "вставить" отложенную команду в существующую последовательность команд. Это позволяет уменьшить количество отдаваемых команд и сделать управление роботом более удобным для оператора.

Несмотря на использование в экспериментах отдельного компьютера и интернет-сервера обработки речи, система легко может быть перенесена на борт робота с использованием одноплатного компьютера и платы обработки речевых команд.

Благодарности. Авторы выражают благодарность Карпову В.Э. за участие в обсуждении данной работы.

Список литературы

- [**Карпов и др., 2000**] Карпов В.Э., Карпова И.П. Язык описания системы контроля знаний // Компьютеры в учебном процессе. 2000. № 4.
- [**Карпова, 2001**] Карпова И.П. Анализ ответов обучаемого в автоматизированных обучающих системах // Информационные технологии. 2001, № 11.
- [**Московский и др., 2018**] Московский А.Д., Бургов Е.В., Овсянникова Е.Е. Зрительный анализатор анимата как основа семантики сенсорной системы робота // Мехатроника, автоматизация, управление. 2018. № 5.
- [**Хачумов, 2012**] Хачумов М.В. Расстояния, метрики и кластерный анализ // Искусственный интеллект и принятие решений. 2012. № 1.
- [**Ющенко, 2009**] Ющенко А.С. Диалоговое управление роботами с использованием нечетких моделей // Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сборник трудов V-й Международной научно-практической конференции. Москва: Физматлит, 2009.
- [**Alonso-Martín et al., 2013**] Alonso-Martín F., Malfaz M., Sequeira J. et al. A multimodal emotion detection system during human-robot interaction // Sensors (Switzerland). 2013. vol. 13. no 11.
- [**Gadanho et al., 2003**] Gadanho S.C. and Custódio L. Learning behavior-selection in a multi-goal robot task // Inform. 2003. vol. 27. no 2.
- [**Ondáš et al., 2013**] Ondáš S., Juhár J., Pleva M. et al. Speech Technologies for Advanced Applications in Service Robotics // 2013. vol. 10. no 5.
- [**Quigley et al., 2009**] Quigley M., Conley K., Gerkey B. et al. ROS: an open-source

- Robot Operating System // ICRA workshop on open source software. 2009.
- [Schuller et al., 2008] Schuller B., Rigoll G., Can S., Feussner H. Emotion sensitive speech control for human-robot interaction in minimal invasive surgery // Proc. 17th IEEE Int. Symp. Robot Hum. Interact. Commun. RO-MAN. 2008, September.
- [Sivakumar et al., 2013] Sivakumar M.S., Murji J., Jacob L.D. et al. Speech controlled automatic wheelchair // 2013 Pan African Int. Conf. Inf. Sci. Comput. Telecommun. PACT 2013, July.
- [Tausczik et al., 2010] Tausczik Y.R. and Pennebaker J.W. The psychological meaning of words: LIWC and computerized text analysis methods // J. Lang. Soc. Psychol. 2010. vol. 29. no 1.

ABOUT APPLYING SENTIMENT ANALYSIS TO VOICE CONTROL IN ROBOTICS

Irina P. Karpova (*karpova@mail.ru*)
National Research University Higher School of Economics,
101000, 20 Myasnitskaya ulitsa, Moscow
National Research Center "Kurchatov Institute", Moscow

Maxim A. Rovbo (*rovboma@gmail.com*)
National Research Center "Kurchatov Institute",
123182, 1 Akademika Kurchatova pl., Moscow

The paper considers the problem of controlling a robot using a voice interface with speech recognition and analysis of the resulting set of words. The proposed method of command recognition is based on a dictionary of commands and special modifier words that are used for sentiment analysis of the command phrase and determining the priority of the task execution.

Keywords: mobile robot, voice control, sentiment analysis