

USING SPEECH RECOGNITION SYSTEM TO CONTROL BUSINESS SYSTEMS AND APPLIANCES

Hasan Gyulyustan

Abstract. Speech recognition is one of the most up-to-date problems that scientists are working about. Actual state of the problem allows to use it for many different applications. The paper represents application of a speech recognition system on a Raspberry Pi based home automation using free web services and software. The software includes combination of different programming and scripting languages for Debian distribution of Linux based operating systems.

Keywords: pi, Raspberry, Raspbian, recognition, speech, system

1. Introduction

Voice or speech recognition is the ability of a machine or program to receive and interpret dictation, or to understand and carry out spoken commands.

For use with computers, analog audio must be converted into digital signals. This requires analog-to-digital conversion. For a computer to decipher the signal, it must have a digital database, or vocabulary, of words or syllables, and a speedy means of comparing this data with signals. The speech patterns are stored on the hard drive and loaded into memory when the program is run. A comparator checks these stored patterns against the output of the A/D converter.

Both acoustic modeling and language modeling are important parts of modern statistically-based speech recognition algorithms. Hidden Markov models (HMMs) are widely used in many systems. Language modeling is also used in many other natural language processing applications such as document classification or statistical machine translation.

Business ICs listed in [1], [2], [3], [4], [5] are good for managing with voice. It can be used to help import data or generate reports.

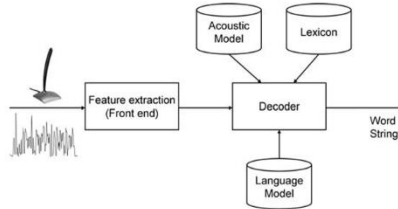


Figure 1. Basic scheme of a typical speech recognition system

2. Design of the recognition system

Hardware for the experiment are raspberry pi boards and a working microphone. Because of that Raspberry Pi doesn't include microphone could be external sound card based microphone, USB microphone or a USB web camera with included microphone, ESP 8266 Wi-Fi microchip and Wi-Fi microchip with infrared.

The Raspberry Pi is a series of small single-board computers developed in the United Kingdom by the Raspberry Pi Foundation to promote the teaching of basic computer science in schools and in developing countries. The original model became far more popular than anticipated, [6] selling outside of its target market for uses such as robotics. Peripherals (including keyboards, mice and cases) are not included with the Raspberry Pi. Some accessories however have been included in several official and unofficial bundles.

The Foundation provides Raspbian, a Debian-based Linux distribution for download, as well as third party Ubuntu, Windows 10 IOT Core, RISC OS, and many specialized media center distributions. It promotes Python and Scratch as the main programming languages, with support for many other languages. The default firmware is closed source, while an unofficial open source is available.

The Raspberry Pi 3 Model B features a quad-core 64-bit ARM Cortex A53 clocked at 1.2 GHz. This puts the Pi 3 roughly 50% faster than the Pi 2. Compared to the Pi 2, the RAM remains the same – 1GB of LPDDR2-900 SDRAM, and the graphics capabilities, provided by the VideoCore IV GPU, are the same as they ever were. As the leaked FCC docs will tell you, the Pi 3 now includes on-board 802.11n WiFi and Bluetooth 4.0. WiFi, wireless keyboards, and wireless mice now work out of the box.

Google Cloud Speech API enables developers to convert audio to text by applying powerful neural network models in an easy to use API. The API recognizes

over 80 languages and variants, to support your global user base. You can transcribe the text of users dictating to an application's microphone, enable command-and-control through voice, or transcribe audio files, among many other use cases. Recognize audio uploaded in the request, and integrate with your audio storage on Google Cloud Storage, by using the same technology Google uses to power its own products.

Google Voice Search or Search by Voice is a Google product that allows users to use Google Search by speaking on a mobile phone or computer, i.e. have the device search for data upon entering information on what to search into the device by speaking

The ESP8266 is a low-cost Wi-Fi microchip with full TCP/IP stack and microcontroller capability. The chip first came to the attention of western makers in August 2014 with the ESP-01 module. This small module allows microcontrollers to connect to a Wi-Fi network and make simple TCP/IP connections using Hayes-style commands. The very low price and the fact that there were very few external components on the module which suggested that it could eventually be very inexpensive in volume, attracted to explore the module, chip, and the software on it.



Figure 2. ESP 8266 chip with relay

Another hardware used for the development of another application of the system is similar to ESP8266 Wi-Fi microchip with infrared module. It could be programmed to control different devices.

3. Development

The software of voice recognition system is developed using C++ programming language.

Figure 3 is a basic scheme which represents the working process of the system.

There are several bash scripts which install and uninstall the software to the system.

The development of this provides to run an action based on the results. Using additional hardware and sensors the following system could be useful to develop more complex systems.

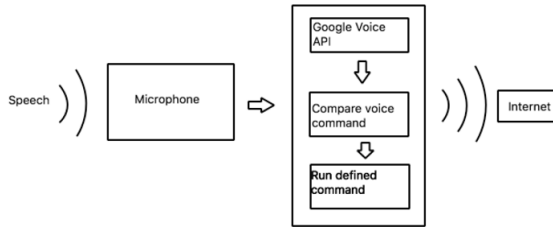


Figure 3. Working scheme of Google speech api based speech recognition software

The following applications of speech recognition are reviewed:

Home automation – Home automation is the automatic or semi-automatic control and monitoring of household appliances and residential house features like doors, gate and even the windows. It is possible to design and build a multi-purpose wireless system that can switch off and on any electrical household appliance, set up them and change settings.

In education – in language learning, speech recognition can be useful for learning a second language. It can teach proper pronunciation, in addition to helping a person develop fluency with their speaking skills.

For text input – Most of the operating systems and websites now allows typing text using speech recognition systems.

In the security systems applications – Today voice biometric technology for authentication user is enough convenient and accurate. This is because the biometrics characteristic if an individual are unique and belongs to the personal until the user dead.

In medical industry – In the health care sector, speech recognition can be implemented in front-end or back-end of the medical documentation process. Front-end speech recognition is where the provider dictates into a speech-recognition engine, the recognized words are displayed as they are spoken, and the dictator is responsible for editing and signing off on the document. Back-end or deferred speech recognition is where the provider dictates into a digital dictation system, the voice is routed through a speech-recognition machine and the recognized draft document is routed along with the original voice file to the editor, where the draft is edited and report finalized. Deferred speech recognition is widely used in the industry currently.

The main application of speech recognitions systems which is a part of following paper is application of home automation systems. It is based on speech detection system developed before using Raspberry Pi board, ESP 8266 Wi-Fi chip and an another infrared chip.

After voice command is detected with speech recognition system using Google's speech API software in Raspberry Pi board communicates with ESP 8266 chip using its API and sends commands to turn on or off relay, change the position of another attached sensor. It's important all of the devices to be in same private network to avoid communication issues.

4. Conclusion

This paper has a research of some applications of public API based speech recognition system with Raspberry Pi as hardware and several programming languages like C and bash. It gathers references to base speech recognition algorithms and structures and their important application in existing speech recognition systems. The paper also covers the hardware specification of the speech recognition system which is very important as size and supported platform and application in home automation.

It maps an application of speech application systems and show a realization using one of them and appropriate hardware and microchips.

Acknowledgements

This work is partially supported by the project FP17-FMI-008 of Faculty of Mathematics and Informatics of Plovdiv University "Paisii Hilendarski".

References

- [1] Valchanov, N., T. Terzieva, V. Shkurtov, A. Iliev, Architecture of extensible computation driven systems, Mathematics and mathematical education, *Proc. of 39th spring conference of Union of Bulgarian Mathematicians*, 06–10 April 2010, Albena, Bulgaria, 2010, 207–211.
- [2] Iliev, A., G. Hristozov, T. Terzieva, Software environment for dynamic models presentation with statistics, *National Conference "Education in the Information Society"*, Plovdiv, 2006, 38–43.
- [3] Valchanov, N., A. Iliev, Implementation of graphical simulation environment for mathematical models, *Fundamental and Complementary Science*, "Mirceacel Batran" Naval Academy Scientific Bulletin, Constanta, Romania, Vol. XIV (2), 2011, 222–228.

- [4] Valchanov, N., P. Petkova, A. Iliev, Integration of computational library for simulation mathematical models in web-based system for courses management, *Proceedings of the Jubilee National Conference with international participation "Tradition, directions, challenges"*, Paisii Hilendarski University of Plovdiv – branch Smolyan, 50 years Scientific and Educational Institution in the Rhodope Mountains, 19–21.10.2012, 2012, 89–94.
- [5] Valchanov, N., A. Iliev, Workflow Optimization using Intelligent Business Information Systems, *Proceedings of the International Conference "Systems Management Business in Small and Medium Enterprises"*, Svishtov, 23-24 April 2010, 107–112.
- [6] Deng, L.; Li, Xiao, (2013), Machine Learning Paradigms for Speech Recognition: An Overview, *IEEE Transactions on Audio, Speech, and Language Processing*, Schmidhuber, Jürgen (2015), Deep Learning, Scholarpedia, 10 (11).
- [7] Oberteuffer, J., Commercial applications of speech interface technology: An industry at the threshold, Proc. Natl. Acad. Sci. USA Vol. 92, 10007–10010, October 1995 Colloquium Paper.
- [8] Deng L., Li, J., Huang, J., Yao, K., Yu, D., Seide, F. et al., Recent Advances in Deep Learning for Speech Research at Microsoft, ICASSP, 2013.
- [9] Bahdanau, Dzmitry, End-to-End Attention-based Large Vocabulary Speech Recognition, 2016.

Факултет по математика и информатика
Пловдивски университет „Паисий Хилендарски“
Бул. „България“ № 236, Пловдив 4003, България
E-mail: hasan@uni-plovdiv.bg

ПРИЛОЖЕНИЕ НА СИСТЕМА ЗА РАЗПОЗНАВАНЕ НА ГЛАС В КОНТРОЛИРАНЕ НА ДРУГИ УСТРОЙСТВА

Хасан Гюлюстан

Резюме. Разпознаването на реч е една от най-актуалните теми, с които учените се занимават напоследък. Досегашните разработки позволяват то да се използва за най-различни приложения. Статията описва приложения на система за разпознаване на глас, базирана на компютър Raspberry Pi в системи като умна къща, използвайки безплатни уеб базирани услуги и софтуер. Софтуерът включва комбинация от различни скриптови и програмни езици за Debian дистрибуцията на операционната система Linux.