

# Voice Command Execution with Speech Recognition and Synthesize

R. Elankavi, Dr.R. Udayakumar

Received: 04 December 2016 • Revised: 07 January 2017 • Accepted: 06 February 2017

**Abstract:** Speech constitutes the primary form of human communication. Speech has the potential to be a better interface other than keyboard and pointing devices. A speech interface would support many valuable applications, like telephone directory assistance, hands busy applications in medicine, office dictation devices, etc. Speech technology is one of the fastest growing modern engineering technology with a wide scope for application in various arenas and disciplines of life. It has many potential benefits and is useful to people in many walks of life. The speech recognition systems in those particular cases provide a significant help to them, so that they can share information with people by operating computer through voice input. This project is designed and developed keeping that factor into mind, and a little effort is made to achieve this aim. This proposed project is capable to recognize the speech and convert the input audio into text. The initial level effort is made to provide help for basic operations as discussed above, but the software can further be updated and enhanced in order to cover more operations.

**Keywords:** Speech Recognition, Assistive Technology, Human Computer Interaction, Spoken-dialog Systems.

## INTRODUCTION

The graphical user interface (GUI) was a major makeover to character user interfaces. The character user interfaces made a very unfriendly user-interface with users having to remember each command to complete any specific task. This restricted the use of computers to largely the scientific and industrial community. But, with the advent of GUI, the computing scenario transformed to being user-friendly, widely applied and application-rich platform. Though many seem to be satisfied with the present GUI interface, we believe that there is still scope for development in the human-computer interface with many an unexplored and less-explored areas. One such less utilized interface is the speech recognition and synthesis. Speech recognition is the technology involving understanding human speech and converting them to a transcript. Speech synthesizer converts a given text into spoken language form. Through this paper, we have tried an application design approach where speech recognition and synthesis is largely the only interface used to perform tasks such as file management, sending and receiving of e-mails, media player and web search. Limited functionalities at experimental stages was done for each of the above tasks. This paper focus primarily on architecture of the system along with the user-application design and issues. The basic principle used in this system is to closely integrate the speech recognition and synthesis (SRS) systems with the underlying operating system. Some interesting applications of developing on such an approach is that this system could be built into an operating system and installed on various embedded devices starting from vacuum cleaner to the audio system of the car to the industrial machines apart from being used in personal computers, this would give the device support for speech recognition and synthesis. In the present form of the project, these systems can be used for self-help information kiosks, reservation systems et al.

## RELATED WORK

At present some popular speech recognition technologies are Google now (Google Inc.), Microsoft speech recognition (Microsoft), Siri (Apple Inc.). These technologies are assistive technologies which provide a variety of applications in their own domain i.e. Google now and Siri are for smartphone

---

R. Elankavi, Assistant Professor, Department of Computer Science and Engineering, BIST, BIHER, Bharath Institute of Higher, Education & Research, Selaiyur, Chennai. E-mail: kavirajcse@gmail.com

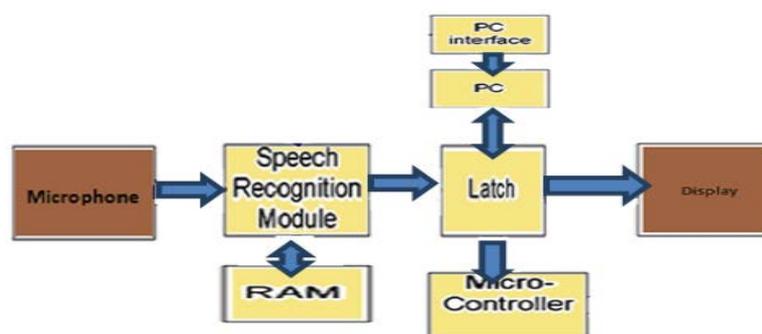
Dr.R. Udayakumar, Professor, Department of IT, BIST, BIHER, Bharath Institute of Higher, Education & Research, Selaiyur, Chennai. E-mail: rsukumar2007@gmail.com

platforms to perform tasks like reminders, check messages and play music. In contrast, Microsoft speech recognition helps to control your computer and dictate/edit text. Microsoft has an inbuilt list of commands which is invoked to control the computer like “click <filename>” or “double click <application>”, these commands simulate the real world computer actions like clicking on mouse and pressing the arrow or scroll key. Apart from that WSR has a few intuitive commands like “open <application name>”, “switch to<application name>” which are more like natural language based control. But, such an approach by Microsoft, though easy to implement and lesser the need for a complete makeover, is entangled in the mess of adapting speech with existing GUI.

This is very noticeable in a feature which allows the following: “Applications that do not present obvious commands can still be controlled by asking the system to overlay numbers on top of interface elements; the number can subsequently be spoken to activate that function. Programs needing mouse clicks in arbitrary locations can also be controlled through speech; when asked to do so, a ‘mouse grid’ of nine zones is displayed, with numbers inside each. The user speaks the number, and another grid of nine zones is placed inside the chosen zone. This continues until the interface element to be clicked is within the chosen zone”. Therefore, such an interface to computer requires patience, user training, and speech recognition training (WSR being speaker dependent). In recent development Microsoft has released a personal assistant called Cortana for computers and smartphones running Windows 10 operating system. Google Now does not allow users to control the computer or mobile, instead it makes use of the Google services like Gmail and Google calendar to extract useful information about the user and present it to the user when required. Google now, therefore will alert you on your next meeting or pull up the required flight data from the data extracted from the user’s Gmail account. Google Now is therefore an intelligent personal assistant than something to control your computing platform like smart phone or computers. However, Google now provides a very intuitive interface through Google cards where the GUI is adapted to suit the speech recognition technologies than the other way around.

Apple’s Siri is more than just a personal assistant, apart from giving answers to live queries like weather, news using many web services. Siri offers control of the company’s products like the iPhone and iPad through speech interface. Hence with Siri you can place a call, send or read messages and create reminders all in your smartphone. Papers similar to creating such assistive technology using speech, works by overlaying speech on top of the GUI resulting in commands such as “cursor right” and “Down”, emulating the need for point and click approach of GUI in speech interfaces. Our approach eliminates such a requirement and is aimed towards creating speech specific applications. Reference Dil1 discuss the various method to develop speech interfaces for web applications. Speech interfaces can be combined with gesture and pattern analysis algorithms to provide a better user experience for users. Telephone systems have been using speech interfaces to provide weather update voice mail system and social networking.

### SYSTEM ARCHITECTURE



### PROPOSED SYSTEM

To identifying the capability of the above said application and working to build its efficiency further to execute high level voice commands given by the user. This finds great advantage in the now fast moving world where the user has to give only voice commands to finish his job. Consequently this application is expected to reduce the time delay in executing commands with GUI. Answering complex questions. Solving maths equations. Reading social network messages, Conversing, Energy savings efficiency, a voice-controlled human-computer interface has been designed that enables severely handicapped individuals to operate a computer. Hands-free computing is any computer configuration where a user can interface without the use of their hands, an otherwise common requirement of human interface devices such as the mouse and keyboard.

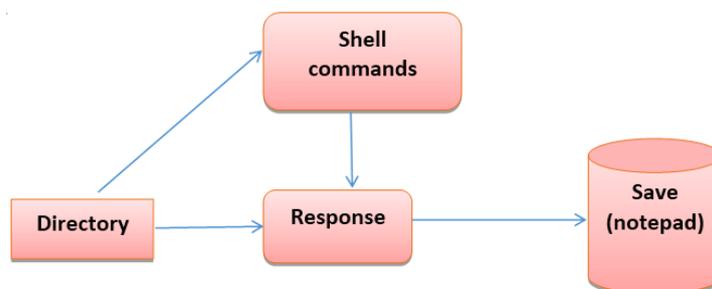
Speech recognition system can be trained to recognize specific commands and upon confirmation of correctness instructions can be given to systems without the use of hands. This may be useful while driving or to an inspector or engineer in a factory environment.

### Admin Login

In this Admin Login module authentication used correct username and password will given to enter into the voice interface recognition providing the business logic to the level various authentication level will be analyzed.

### Shell Commands

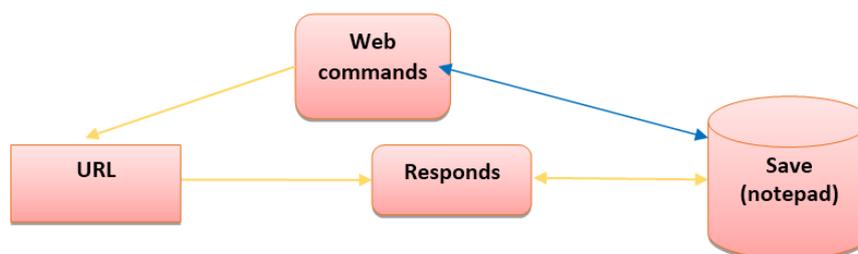
Location of every file, folder and application is to be specified initially. It is recommended to provide related commands. Colloquial languages are difficult to recognize for Microsoft speech recognizer. Any application integration can be done with the help of this module.



### Web Commands

By integrating with the user info, it is used to read the e-mail specified by the user. All web pages can be accessed using your default web browser with the help of this module.

This module also consists of 3 inputs initially. It processes as per the system's firewall and internet security.



### CONCLUSION

The Proposed System of speech recognition started with a brief introduction of the technology and its applications in different sectors. The project part of the Report was based on software development for speech recognition at the later stage we discussed different tools for bringing that idea into practical work

### FUTURE ENHANCEMENTS

The future versions of the system can be expanded to address the following shortcomings

- Being an online system primarily, network lag gave delay in returning results for speech. A better offline online combination can help to compensate for the network lags.
- Speech recognition in noisy environments produces inaccurate results. Noise cancellation algorithms built into the system and better hardware can improve results.
- Error correction for verbal-input is difficult. A button can be included in the system to re-read wrongly inputted speech or wrongly recognized commands.

The system built at a larger scale can include the following.

- Visual elements to complement speech interface
- Support for running multiple applications
- Novel file manager designed for speech interfaces
- Understanding the content in the files would help in retrieving and locating files quickly in speech interfaces. Multimedia files can also be retrieved as discussed in [1].

## REFERENCES

- [1] Das, J., Das, M.P., & Velusamy, P. (2013). Sesbania grandiflora leaf extract mediated green synthesis of antibacterial silver nanoparticles against selected human pathogens. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 104, 265-270.
- [2] Umanath, K.P.S.S.K., Palanikumar, K., & Selvamani, S.T. (2013). Analysis of dry sliding wear behaviour of Al6061/SiC/Al2O3 hybrid metal matrix composites. *Composites Part B: Engineering*, 53, 159-168.
- [3] Udayakumar, R., Khanaa, V., Saravanan, T., & Saritha, G. (1786). Cross layer optimization for wireless network (WIMAX). *Middle-East Journal of Scientific Research*, 16(12), 1786-1789.
- [4] Kumaravel, A., & Rangarajan, K. (2013). Algorithm for automaton specification for exploring dynamic labyrinths. *Indian Journal of Science and Technology*, 6(5S), 4554-4559.
- [5] Pieger, S., Salman, A., & Bidra, A.S. (2014). Clinical outcomes of lithium disilicate single crowns and partial fixed dental prostheses: a systematic review. *The Journal of prosthetic dentistry*, 112(1), 22-30.
- [6] Vijayaraghavan, K., Nalini, S.K., Prakash, N.U., & Madhankumar, D. (2012). One step green synthesis of silver nano/microparticles using extracts of Trachyspermum ammi and Papaver somniferum. *Colloids and Surfaces B: Biointerfaces*, 94, 114-117.
- [7] Khanaa, V., Mohanta, K., & Satheesh, B. (2013). Comparative study of uwb communications over fiber using direct and external modulations. *Indian Journal of Science and Technology*, 6(6), 4845-4847.
- [8] Khanaa, V., Thooyamani, K.P., & Udayakumar, R. (1798). Cognitive radio based network for ISM band real time embedded system. *Middle-East Journal of Scientific Research*, 16(12), 1798-1800.
- [9] Vijayaraghavan, K., Nalini, S.K., Prakash, N.U., & Madhankumar, D. (2012). Biomimetic synthesis of silver nanoparticles by aqueous extract of Syzygium aromaticum. *Materials Letters*, 75, 33-35
- [10] Caroline, M.L., Sankar, R., Indirani, R.M., & Vasudevan, S. (2009). Growth, optical, thermal and dielectric studies of an amino acid organic nonlinear optical material: l-Alanine. *Materials Chemistry and Physics*, 114(1), 490-494.
- [11] Kumaravel, A., & Pradeepa, R. (2013). Efficient molecule reduction for drug design by intelligent search methods. *International Journal of Pharma and Bio Sciences*, 4(2), B1023-B1029.
- [12] Kaviyarasu, K., Manikandan, E., Kennedy, J., Jayachandran, M., Ladchumananandasivam, R., De Gomes, U.U., & Maaza, M. (2016). Synthesis and characterization studies of NiO nanorods for enhancing solar cell efficiency using photon upconversion materials. *Ceramics International*, 42(7), 8385-8394.
- [13] Sengottuvel, P., Satishkumar, S., & Dinakaran, D. (2013). Optimization of multiple characteristics of EDM parameters based on desirability approach and fuzzy modeling. *Procedia Engineering*, 64, 1069-1078.
- [14] Anbuselvi S., Chellaram, C., Jonesh S., Jayanthi L., & Edward J.K.P. (2009). Bioactive potential of coral associated gastropod, Trochus tentorium of Gulf of Mannar, Southeastern India. *J. Med. Sci*, 9(5), 240-244.
- [15] Kaviyarasu, K., Ayeshamariam, A., Manikandan, E., Kennedy, J., Ladchumananandasivam, R., Gomes, U.U., & Maaza, M. (2016). Solution processing of CuSe quantum dots: Photocatalytic activity under RhB for UV and visible-light solar irradiation. *Materials Science and Engineering: B*, 210, 1-9.
- [16] Kumaravel, A., & Udayakumar, R. (2013). Web portal visits patterns predicted by intuitionistic fuzzy approach. *Indian Journal of Science and Technology*, 6(5S), 4549-4553.
- [17] Srinivasan, V., & Saravanan, T. (2013). Reformation and market design of power sector. *Middle-East Journal of Scientific Research*, 16(12), 1763-1767.
- [18] Kaviyarasu, K., Manikandan, E., Kennedy, J., & Maaza, M. (2015). A comparative study on the morphological features of highly ordered MgO: AgO nanocube arrays prepared via a hydrothermal method. *RSC Advances*, 5(100), 82421-82428.
- [19] Kumaravel, A., & Udhayakumarapandian, D. (2013). Construction of meta classifiers for apple scab infections. *International Journal of Pharma and Bio Sciences*, 4(4), B1207-B1213.
- [20] Sankari, S.L., Masthan, K.M.K., Babu, N.A., Bhattacharjee, T., & Elumalai, M. (2012). Apoptosis in cancer-an update. *Asian Pacific journal of cancer prevention*, 13(10), 4873-4878.

- [21] Harish, B.N., & Menezes, G.A. (2011). Antimicrobial resistance in typhoidal salmonellae. *Indian journal of medical microbiology*, 29(3), 223-229.
- [22] Manikandan, A., Manikandan, E., Meenatchi, B., Vadivel, S., Jaganathan, S.K., Ladchumananandasivam, R., & Aanand, J.S. (2017). Rare earth element (REE) lanthanum doped zinc oxide (La: ZnO) nanomaterials: synthesis structural optical and antibacterial studies. *Journal of Alloys and Compounds*, 723, 1155-1161.
- [23] Caroline, M.L., & Vasudevan, S. (2008). Growth and characterization of an organic nonlinear optical material: L-alanine alaninium nitrate. *Materials Letters*, 62(15), 2245-2248.
- [24] Saravanan T., Srinivasan V., Udayakumar R. (2013). A approach for visualization of atherosclerosis in coronary artery. *Middle - East Journal of Scientific Research*, 18(12), 1713-1717.
- [25] Poongothai, S., Ilavarasan, R., & Karrunakaran, C.M. (2010). Simultaneous and accurate determination of vitamins B1, B6, B12 and alpha-lipoic acid in multivitamin capsule by reverse-phase high performance liquid chromatographic method. *International Journal of Pharmacy and Pharmaceutical Sciences*, 2(4), 133-139.
- [26] Udayakumar, R., Khanaa, V., & Saravanan, T. (2013). Synthesis and structural characterization of thin films of SnO<sub>2</sub> prepared by spray pyrolysis technique. *Indian Journal of Science and Technology*, 6(6), 4754-4757.
- [27] Anbazhagan, R., Satheesh, B., & Gopalakrishnan, K. (2013). Mathematical modeling and simulation of modern cars in the role of stability analysis. *Indian Journal of Science and Technology*, 6(5S), 4633-4641.
- [28] Caroline, M.L., & Vasudevan, S. (2009). Growth and characterization of bis thiourea cadmium iodide: A semiorganic single crystal. *Materials Chemistry and Physics*, 113(2-3), 670-674.
- [29] Sharmila, S., Jeyanthi Rebecca, L., & Das, M.P. (2012). Production of Biodiesel from *Chaetomorpha antennina* and *Gracilaria corticata*. *Journal of Chemical and Pharmaceutical Research*, 4(11), 4870-4874.
- [30] Thooyamani, K.P., Khanaa, V., & Udayakumar, R. (2013). An integrated agent system for e-mail coordination using jade. *Indian Journal of Science and Technology*, 6(6), 4758-4761.
- [31] Caroline, M.L., Kandasamy, A., Mohan, R., & Vasudevan, S. (2009). Growth and characterization of dichlorobis l-proline Zn (II): A semiorganic nonlinear optical single crystal. *Journal of Crystal Growth*, 311(4), 1161-1165.
- [32] Caroline, M.L., & Vasudevan, S. (2009). Growth and characterization of L-phenylalanine nitric acid, a new organic nonlinear optical material. *Materials Letters*, 63(1), 41-44.
- [33] Kaviyarasu, K., Xolile Fuku, Gene T. Mola, E. Manikandan, J. Kennedy, and M. Maaza. Photoluminescence of well-aligned ZnO doped CeO<sub>2</sub> nanoplatelets by a solvothermal route. *Materials Letters*, 183(2016), 351-354.
- [34] Saravanan, T., & Saritha, G. (2013). Buck converter with a variable number of predictive current distributing method. *Indian Journal of Science and Technology*, 6(5S), 4583-4588.
- [35] Parthasarathy, R., Ilavarasan, R., & Karrunakaran, C. M. (2009). Antidiabetic activity of *Thespesia Populnea* bark and leaf extract against streptozotocin induced diabetic rats. *International Journal of PharmTech Research*, 1(4), 1069-1072.
- [36] Gupta, T., & Sharma, A. (2014). Search Accuracy in Web information Retrieval. *International Journal of Communication and Computer Technologies*, 2(2), 98-105.
- [37] Kakavand, Z.M., & Chalechale, A. (2015). Comparison of Two Different Distance Functions of Image Retrieval for Detecting Species of Microscopic Fungi in Medical Mycology Laboratory. *International Academic Journal of Science and Engineering*, 2(4), 39-44.
- [38] Ramya, V., Ranjitha, S., Sathya Sofia, A., & Ganesh Kumar, P. (2014). Load Balancing of Tasks in Cloud Computing Environment Using Honey Bee Behavior. *International Journal of System Design and Information Processing*, 2(2), 25-53.
- [39] Brindha, M.S. (2017). A Survey on Cross Layer Distributed Topology Control in Mobile Adhoc Network. *Bonfring International Journal of Networking Technologies and Applications*, 4(1), 1-3.
- [40] Kumar, P. (2014). Load Characteristics of Electric System for Distributing Power on Locality Based Criterion. *Bonfring International Journal of Power Systems and Integrated Circuits*, 4(4), 39-43.

- [41] Rama Rao, G., Purna Prakash, J., & Rama Raju, M. (2014). Designing High to Low Cost Solution for Crash Recovery of Servers through Virtualization using Cloud Computing. *International Scientific Journal on Science Engineering & Technology*, 17(5), 549-555.
- [42] Arul Jothy, K., Sivakumar, K., & Delsey, M.J. (2018). Distributed System Framework for Mobile Cloud Computing. *Bonfring International Journal of Research in Communication Engineering*, 8(1), 5-9.
- [43] Kabeer, V., & Zainul Abid, T.P. (2013). Automated Face Recognition using Artificial Light Receptor Model and SVM Classifier. *The SIJ Transactions on Computer Science Engineering & its Applications*, 1(3), 36-41.
- [44] Yasvanthkumaar, V., Sabitha, S., & NithyaKalyani, S. (2018). Parallel and Multiple E-Data Distributed Process with Progressive Duplicate Detection Model. *Bonfring International Journal of Software Engineering and Soft Computing*, 8(1), 23-25.
- [45] Kanaga Sundar, R., Joe Paul, J., Krishna Kumar, M., & Laser, L. (2014). Secured POR for Flooding Attack Prevention in Extremely Dynamic Ad Hoc Networks. *The SIJ Transactions on Computer Networks & Communication Engineering (CNCE)*, 2(2), 5-9.
- [46] Hanirex, D.K., & Kaliyamurthie, K. P. (2013). Multi-classification approach for detecting thyroid attacks. *International Journal of Pharma and Bio Sciences*, 4(3), B1246-B1251.
- [47] Kandasamy, A., Mohan, R., Lydia Caroline, M., & Vasudevan, S. (2008). Nucleation kinetics, growth, solubility and dielectric studies of L-proline cadmium chloride monohydrate semi organic nonlinear optical single crystal. *Crystal Research and Technology: Journal of Experimental and Industrial Crystallography*, 43(2), 186-192.
- [48] Srinivasan, V., Saravanan, T., Udayakumar, R., & Saritha, G. (2013). Specific absorption rate in the cell phone user's head. *Middle-East Journal of Scientific Research*, 16(12), 1748-50.
- [49] Udayakumar R., Khanaa V., & Saravanan T. (2013). Chromatic dispersion compensation in optical fiber communication system and its simulation. *Indian Journal of Science and Technology*, 6(6), 4762-4766.
- [50] Vijayaragavan, S.P., Karthik, B., Kiran, T.V.U., & Sundar Raj, M. (1990). Robotic surveillance for patient care in hospitals. *Middle-East Journal of Scientific Research*, 16(12), 1820-1824.