IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

Speech Based Interactive System for Physically Challenged People

Prasanna Deshmukh[1], Swapnil Shedge[2], Pranav Dhole [3], S. P. Patil[4]

Department of Information Technology, NBN Sinhgad School of Engineering, Pune, Maharashtra, India[1,2,3]

Professor, Dept. of Information Technology, NBN Sinhgad School of Engineering, Pune, Maharashtra, India[4]

Abstract: In a study about how humans and robots can successfully interact to accomplish specific tasks is important in creating more sophisticated robots that may eventually become a part of human society. Advances in human-computer interaction technologies are being extensively used in improving human–robot interaction (HRI). This paper demonstrates an interactive and always on system that helps people to carry out specific tasks like check Facebook notification, weather, news, etc. by using natural voice commands effortlessly. The main objective of this work is to help physically disabled people to interact with the system without any physical contact. The system uses speech recognition to get the query, processing is done using cloud based third party services and Text-to-Speech (TTS) is used for the results.

keywords: Human-Robot interaction, Text-to-Speech, Speech-to-Text, Cloud Services, Raspberry pi.

I. INTRODUCTION

As mentioned in the survey that was carried out by the World Health Organization (WHO), just over a billion people, about 15% of the world's population, have some form of disability [1]. The counts of adults ranging between 110 and 190 million have notable difficulties in functioning. All these people need to maintain the pace of their lives with the changing technology in the world. What if, a system is developed that will allow all these people to interact with it just by using natural voice commands? It will surely help deal with the difficulties confronted by people having disabilities, who are not able to seek help of technology to carry out their daily tasks.

The single most important objective in developing this system is to help physically challenged people carry out certain tasks without having to actually reach out to the system and query it to obtain the result. The system will be always on and in a sleep mode.

The person will be able to wake the system just by speaking the trigger word and then query the system using voice commands. The foremost advantage of the system is that, the need of user's physical contact with the system is eliminated. The system will respond to the user's speech based commands, process the query using cloud services and convert the resulting text into speech to output the result.

II. RELATED WORK

It has been observed that, several advancements in making this kind of systems are being done. All the systems developed lack some functionalities due to one or other limitations enforced by the way the user can communicate with the system. There are systems that lack accuracy of input and have communication loopholes. This results in poor performance of the system and thereby limiting the functionalities being provided. Currently available systems use high power consuming CPU for processing the queries that lower the throughput of the system. The proposed system uses cloud based speech recognition system [2][3]for increased accuracy in the input. It also uses cloud based infrastructure in order to off-load heavy tasks to the cloud[5][6].

The author Erico Guizzo in [5] has proposed a system that makes the robots that are to be designed to be connected to the clouds. According to the author, if the robots that are eventually an embedded system of various electronic parts if connected to the cloud can perform calculations and various functions quite swiftly and efficiently. He designed a robotic multimodal system that relies on the cloud services to enumerate an environment of the human and robot interaction that results in a brain storming effort towards this technology.

The author in [5] has developed and devised a system that controls a robotic hand that is able to do some heavy work. The robotic hand can be controlled remotely from any location in the world with help of the cloud services provided.

The authors Yan-You Chen, Jhing-Fa Wang in [6] has proposed a human- robot for senior companion based on cloud computing infrastructure.

They have researched various robotic multimodal systems that help in making the life of the few elderly and some disabled people much more comfortable and make them autonomous. In the robot client side, the behaviour model is designed to use WURMS and RMICS services. With

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

using only low-cost and low-power CPUs (Intel Atom N450), both of the two robots can still work wirelessly for real-time human-robot interaction.

After studying the previous work [5][6]it is observed that the previously designed systems used power hungry processors which led to processing inefficiency. Compared to that our system uses very efficient and low power consuming processor that increases the responsiveness of the system.

III. PROPOSED SYSTEM

A. System Introduction

Technology changes as time advances and so does systems based on technology do. Being physically challenged is quite unfortunate and hence leads to dependency on others to perform certain tasks. At times, tasks to be performed either need the user to physically interact with the system or to move from one place to other, just to get the work done. As a result of the disability, the user's movement are limited to certain extent and thereby increases the need to rely on others. This hampers the productivity of the user and reduces their morale.

The proposed system provides a user friendly way of communication i.e. natural voice commands (speech to text)[2] [3] for querying the system and cloud based text to speech [7] is used to narrate the results. The system uses Raspberry Pi board [8] which provides a low cost alternative to processors with high processing powers. A microphone connected to the Pi is used to input the voice commands as entered by the user and external speakers are used for the output. The advantage of the system is that voice commands can be used to control the system rather than actually reaching out to the system for queries. On the other hand, the system is always in an active listening mode that improves the responsiveness of the system. Finally, cloud services are used to synthesize the speech input, process the result and expatiate the output through external speakers.

B. System Architecture

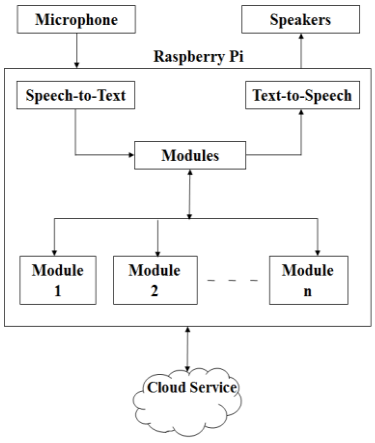

Fig. 1 System Block Diagram

As shown in the diagram [fig 1], [I]an external microphone is connected to the system that captures the input for the speech based query, [II] the speech-to-text engine converts the speech into text format and passes it to the corresponding module, [III] the corresponding module processes the input actively or passes to the cloud services for processing, [IV] the particular cloud service provides the answer back to the calling module, [V] the result in the text format is then passed to the text-to-speech engine(TTS), [VI] output from the TTS is passed to external speakers to expatiate the result.

C. System Modules

There are various modules in the system that perform specific tasks and are described below:

1. Time module provides current clock time with respect to the time zone of the system.
2. Weather module provides the user with current weather conditions and forecasts based on the location ID provided by the user.
3. News module entertains the user with top news articles and can also provision sending an email to the user with the link to detailed news.
4. Gmail module will be used to notify the user about newly arrived emails.

D. Algorithm

Below is the procedural working of the system that explains the working of the modules in a sequential manner:

1. Start
2. System initially is in Active/Sleep mode
3. User uses the trigger word to wake the system up
4. if (Trigger word spoken)
{
System responds with a small beep and enters into the listening mode
}
else
{
System responds with a keyword "Pardon"
goto 2
}

5. if (valid query asked)
{
System calls a relevant module and responds with an appropriate answer by making use of cloud services
if (User input required and given)
{
System responds with a long beep and provides appropriate answer output depending on the given input by the user
}
else
{
System responds with a keyword "Pardon"
goto 2
}
else
{

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

System responds with a keyword "Pardon"
goto 2
}

6. Stop

E. Mathematical Model

Below is the mathematical representation of the system using set theory:

A. Let S be a System in Execution
S = {....}

B. Identify the modules as M
S = {M...}
M = {U}
Where,
U = User
Identify module U
U = {T,W,N,G,Q}
Where,
T = Time Module
W = Weather Module
N = News Module
G = Gmail Module
Q = Quote Module

1. Identify module as U as T
T = {Te}
Where,
Te = Time Expatiate

2. Identify module as U as W
W = {We}
Where,
We = Weather Expatiate

3. Identify modules as U as N
N = {Ne, Ns}
Where,
Ne=Top Trending News Expatiate
Ns = News Send via Email

4. Identify module as U as G
G = {Gc}
Where,
Gc = Expatiate Unread Mail count

C. Identify the Processes as P
S = {M, P...}
P = {Pi,Pr,Po}
Where,
Pi = Processing Input
Po = Processing Output using Third Party Cloud Services
Pe = Output Expatiate

D. Identify the output as O
S = {M,P,O, …}
O = {Ot,Ow,On,Og}
Where,
Ot = Precise Time based on Timezone
Ow = Accurate Weather based on Location ID
On = Top Trending News
Og = Unread Mail Count

E. Identify the Success as Su
S = {M,P,O,Su,...}
Where,
Su = Success is When user gets accurate answer to the asked queries by using third party cloud services

F. Identify the Failure as F
S = {M,P,O,Su,F, …}
Where,
F = Modules fail to execute and behave anomalously.

The System can be described as
S = {M,P,O,Su,F}

IV. SYSTEM FEATURES

1. Always on System
   The system remains always active in a low power state so as to facilitate responsiveness.
2. Quick response to user query
   The system possesses highly sensitive speech recognition for quicker response.
3. Accurate result
   Availability of cloud services makes system smart enough to produce accurate results.
4. Offloading heavy task to the cloud
   Cloud services used by the system helps in reducing the use of processing power and resulting in high throughput.

V. EVALUATION TEST CASES

Below given table shows test to pass evaluation scenario of the system:

Table I

Test to Pass

Test Case No.	Description	Expected Result	Actual Result
1	Testing the system's response to the trigger word.	Small Beep	Small Beep
2	Testing noisy speech input from the user	System responds with a keyword (e.g. "Pardon")	System responds with a keyword (e.g. "Pardon")
3	Testing the system's response without internet connection	Anomalous behaviour of the system with a warning.	System reacts in an unexpected way
4	To test module priority having same trigger keyword	Module with higher priority should respond first	High priority module responds
5	To test responsiveness of each module	Module should fail under certain circumstances	Modules fail
6	Testing the System's response to specific keywords (e.g. News, Time, Weather)	System calls the particular modular	All the modules called

Below given table shows test to fail evaluation scenario of the system:

Table II
Test to Fail

Test Case No.	Description	Expected Result	Actual Result
1	Testing the system's response to other than trigger word.	System should not responds	System ignores the words.
2	Input an irrelevant keyword or phrase to module	module should fail to response	Module fails with a keyword "Pardon"
3	Simultaneous speech input to module	Creation of multiple instances and abrupt behaviour	Module create instances and fails to response

VI. CONCLUSION AND FUTURE WORK

In this paper, we design an "always on" voice control system which uses cloud services. Also being deployed by using Raspberry Pi faster processing is available with less power consumption thereby reducing the overall processing latency. Thus use of this system will bridge the gap between physically challenged people and the technology so that they can perform daily tasks effortlessly and independently.

In the future, the system will be able to identify the user based on either facial recognition or voice signatures. Another addition to the system would be the use of offline speech to text and text to speech engines in order to retain the system's responsiveness and increased privacy. Also, the system will support multiple users with simultaneous input resolution.

REFERENCES

[1]. http://www.who.int/disabilities/world_report/2011/report/en
[2]. http://www.google.com/intl/en/chrome/demos/speeh.html
[3]. http://gist.github.com/alotaiba/1730160
[4]. Fong,T.Thorpe and Baur, "Collaboration dialogue and human-robot interaction", proc. 10th International Symposium on Robotics Research, Springer, 2001.
[5]. Erico Guizzo, "Robots with Their Head in the Clouds," IEEE SPECTRUM, ROBOTICS, MARCH 2011.
[6]. Yan-You Chen, Jhing-Fa Wang "Human-Robot Interaction Based on Cloud Computing Infrastructure for Senior Companion" by National Cheng Kung University, Tainan, Taiwan, TENCON 2011.
[7]. http://www.ivona.com
[8]. http://www.raspberrypi.org/wp-contetnt/uploads/2012/12/quick-start-guide-v1.1.pdf

DOI 10.17148/IJARCCE.2016.5336