CASOR: Context-Aware Social Robot for Home Environment

[1]Ankur Pandey, [2]Rupali Nikhare

[1]M.E. Student, [2]Assistant Professor

[1]Information Technology Department(AI & R),Pillai College of Engineering, New Panvel, Mumbai, Maharashtra, India

[2] Computer Engineering Department, Pillai College of Engineering, New Panvel, Mumbai, Maharashtra, India

***Abstract* - Parallel research in artificial intelligence and robotics has created a new outlook towards the capabilities of machines. The era of robotics research started from simple mechanical industrial machines to complex intelligent robots that can learn from experience and mimic humans in everyday tasks. Social robots are one such example of artificial intelligent robots. Robots are becoming an integral component of our society and have great potential in being utilized in a variety of ways to serve humans. The existing systems are way too complex to study, recreate, less flexible and cost ineffective. A low cost social robotic platform is presented in this paper. The robot is context aware in home environment and is comprised of various sensors that can monitor temperature, humidity and light intensity in home conditions. It also consists of passive infrared sensor to detect nearby presence of user. The robot can control the switching of home devices with the help of voice commands. The robotic system can be monitored remotely using Internet. It uses open source hardware and software technologies in order to be economical.**

***Index Terms* - Context-aware social robot, Casor social robot, robot home automation, voice controlled robot**

I. INTRODUCTION

The main objective of the social robotics research is to design robots that involve in social scenarios which are familiar and compelling to humans. Thus, the robots have to provide a social communicative functionality that is intuitive and natural. This can be supported if appearance and functionality fit the robot's tasks and robots are as self-explaining as possible. These points are very important while designing robots to interact with naive users in domestic environments.

As interest in robots and robotics as a research field increases, there is a parallel interest in the social aspects of robotics as well. The goal is to create robots that can work hand in hand with humans and collaborate with them in real world environments. In order to achieve this successfully, such robots need to be context-aware and should be able to interpret and produce a wide range of social functions. This includes principles of design, communication, engineering, sociology and human-robot interaction which makes robotics a multidisciplinary field [2]. Social robot can exit in the human society and it is more likely to get adapted by masses if we consider the development of such robot from user's perspective [1]. Home automation which involves the control and automation of various household devices like light, television, air conditioner etc is adding convenience in human lives and can be integrated with robots to provide such services.

II. LITERATURE SURVEY

In this section we cite the relevant past literature that describes the study, design and implementation of social robots using various techniques. Most of the researchers focused on complex design aspects and mechanical interaction. Multiple sensor system is rarely used in order to receive feedback from surrounding environment.

Dautenhahn, Kerstin et al. [8] designed KASPAR, a minimally-expressive humanoid robot as a therapeutic toy for children with autism.The interaction of this robot is guided by the experimenter. It consisted of a variety of gestures and facial expressions.

Michael Ferguson et al. [5] described a social robotic platform named 'Nelson' for educational and research purpose. The robot consist of a social face, arms for representing gestures, advance sensor system etc. The complete robot is controlled by sending serial commands using a computer.

Sangmeshwar Kendre and S. D. Apte [7] demonstrated the design and implementation of a mobile robot which can be controlled using voice commands. The robot is capable of interpretation of only five words which are forward, reverse, left, right and stop. It uses wavelet denoising and speech recognition using TIESR speech recognizer. It is based on Beaglebone hardware platform.

A low cost home automation system is projected by Prasanna G. and Ramadass N [4]. The hardware implemented uses Hidden Markov Model Toolkit (HTK). Voice commands are converted into text messages using HTK. These messages are transmitted and received using GSM modem. Thus based on text messages, appliances are controlled.

The design of mobile service robot that can be controlled with some predefined voice commands and runs on a fixed path as line following feature is embedded is proposed by Vaishali Wagh and Manisha Wasnik. It has a centralized server which sends commands wirelessly to robot using zigbee technology to identify and reach to home appliances using respective device tags [3].

III. METHODOLOGY

The context-aware social robot is a complex system and has various important building blocks. As the system, we built and focused on the head part of robot with jaw motion and voice as feedback. The robot is capable of processing voice commands and performs home automation. In the Fig.1, overview of the complete system is given.

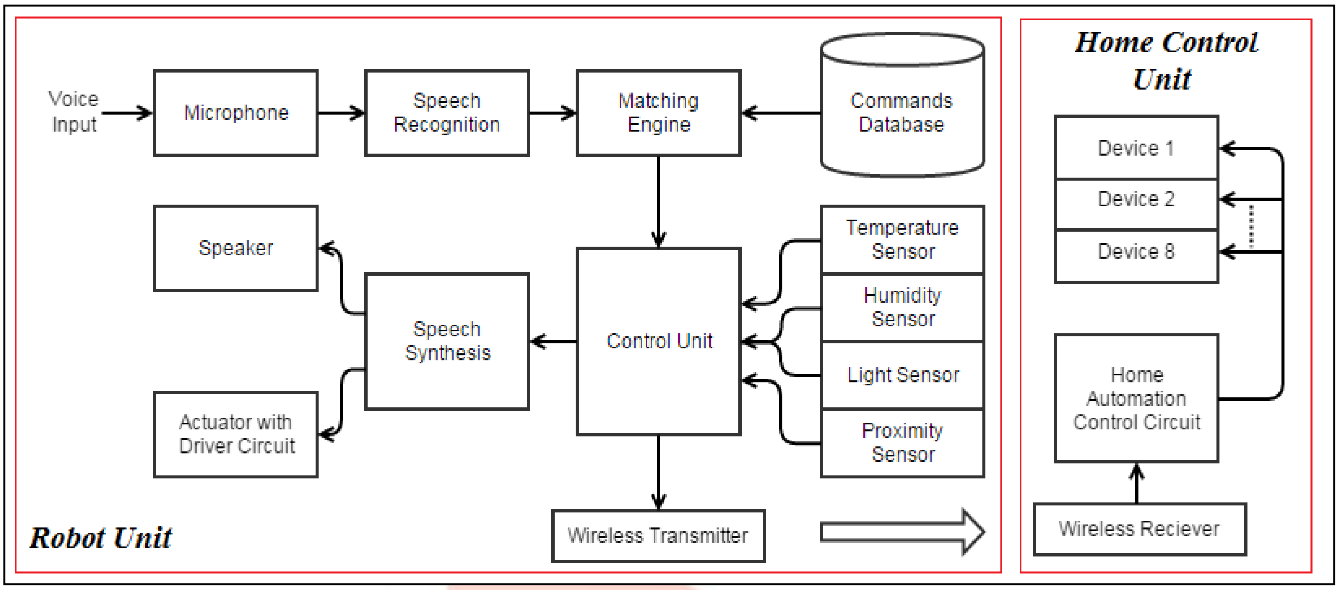


Figure 1 Block diagram of the complete system

The robot has the capability to interact with the user through voice interaction. The robot can accept voice commands as input then the voice command are transferred from the microphone to speech recognition system, in our case PocketSphinx is used as Speech recognition toolkit [10]. After the system converts speech signals to text, a matching engine evaluates the recognized speech with the set of commands present in the commands database. Commands database is basically a mapping of key and value. All the commands have a particular associated action. The mapped key-value pair is passed to control unit which is the main program that is responsible for output generation.

The control unit then provides the output as text to speech synthesis system where text to speech processing is done. We are using eSpeak speech synthesis tool to serve the purpose. Speech synthesis module will provide the output to speaker and to servo driver unit. Synchronization of servo motion and output audio is also implemented.

Robot Unit

The robot unit is a combination of mechanical parts, actuators and other electronics which makes the head part of the robot. The robot is controlled using raspberry pi which acts as processing unit. Since raspberry pi is powered by open source software and runs a complete debian linux version along with good number of input /output pins, it was sufficient as per the requirement of the project. Raspberry pi is used by a large number of people worldwide and has a strong community support. These were the reasons to choose raspberry pi as the main processing unit for robot unit. It also consists of servo motors as actuators for controlling head and jaw movements of robot and sensors to achieve context awareness in home environment. Microphone and speaker will act for voice input and output respectively. For wireless communication there is wifi module and radio frequency transmitter.

Home Control Unit

The home control unit consists of radio frequency receiver to receive signals from robot unit to switch a particular device. The controlling unit for this module is Arduino Uno R3 development board. This version of arduino is powered by ATmega328 and has 14 General Purpose Input/ Output (GPIO) pins. The home control unit also consists of an 8 channel relay board for the process of home automation.

IV. DESIGN AND DEVELOPMENT

The robotic design used in this project is part of InMoov open source robot [15]. InMoov is the first Open Source 3D printed life-size robot designed by Gael Langevin, who is a French model maker and sculptor. The design of head of this open source robot is used and is further extended by external cosmetic changes. The design of the robotic head which is adapted for the project is shown in Fig. 2.

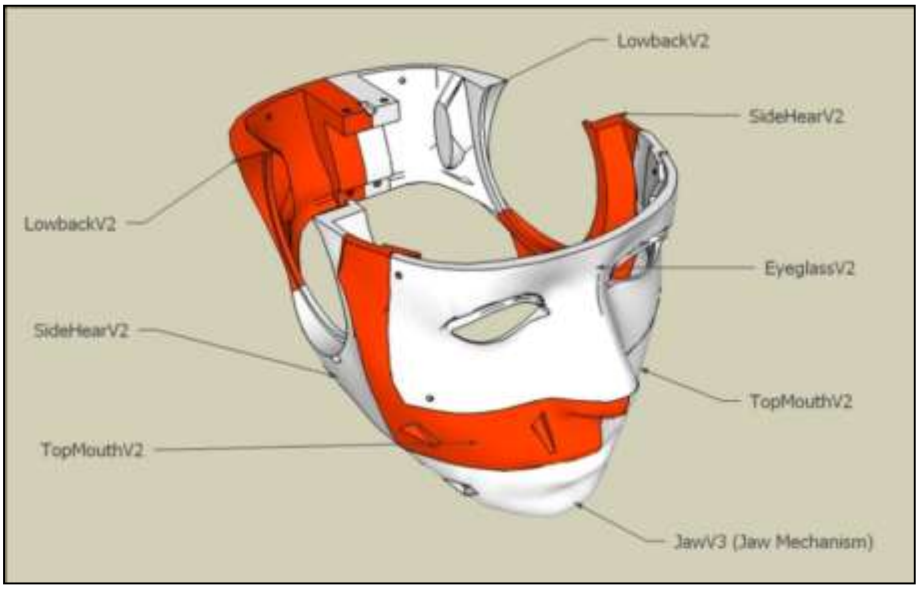


Figure 2 InMoov open source humanoid robot head design [15]

3D Printing Process

3D printing also known as additive manufacturing (AM) is a set of different processes in which three dimensional objects are synthesized. In this process, very fine successive layers of material are laid to create an object. The complete process is under computer control with the help of dedicated software. This software uses 3D digital model of object to create its realistic model by sending instruction to 3D Printer. We have used white Polylactic Acid (PLA) filament in order to print the robot.

Felix 3.1 3D Printer which has a build volume of 25.5 x 20.5 x 23.5 cm and a heated build platform was used to print the different parts of the robotic head. It prints at a maximum resolution of 50 microns. The design files uses STereoLithography (STL) format. An open source software know as Repetier was used to open the 3d file of object and set parameters like printer's bed temperature, extruder's temperature and orientation position of object on build platform. We can add and position the STL files on the simulated print bed and slice them all together. The different individual parts required are 3D printed one by one as shown in Fig.3. In Fig.4 (a) 3D design of eye mask of robot is opened in Repetier software Fig.3 (b) and (c) shows the eye mask of robot getting printed and completed print respectively.

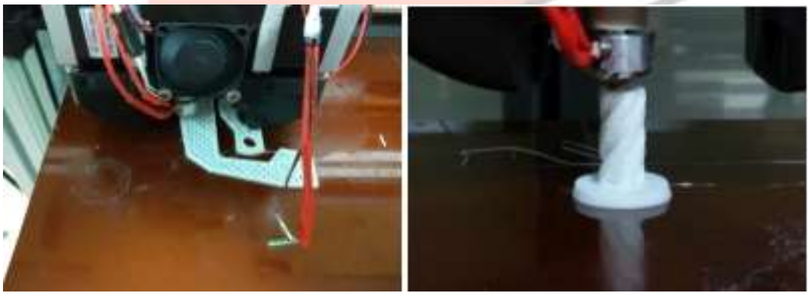
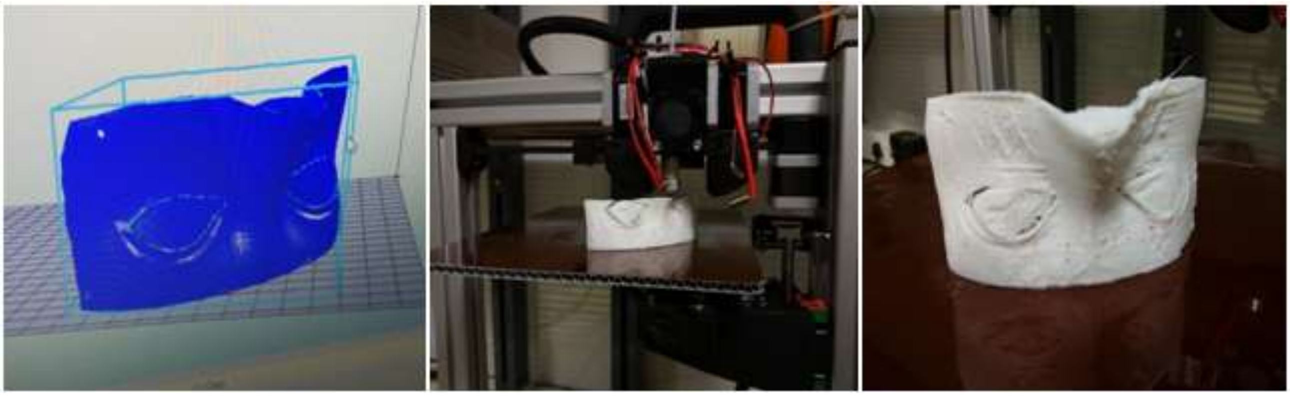


Figure 3 Printing of different parts required for robotic head assembly



(a) (b) (c)

Figure 4 3D Printing of Robotic Head Part (Eye Mask)

One by one all the parts required for the assembly of robotic head were printed. Fig.5 (a) shows the complete set of different parts which are printed and required for robotic head. In order to enhance the appearance of robot, an external plastic mask that mimics human face is added as shown in Fig.5 (b). The completed robotic face after assembling the 3D printed parts and cosmetic enhancement is shown in Fig. 5 (c)

(a) (b) (c)

Figure 5 The different phases while building a robotic head

V. SYSTEM IMPLEMENTATION

The complete robotic system is divided into two parts which are the robot unit and home control unit. The first is the robot unit. In Fig.6 the robot unit which is connected to robotic head is shown. It consists of Raspberry Pi B+ as the main controlling unit. An external audio adapter is connected to it for interfacing microphone for voice input and speaker for output. A Sensor card is attached to raspberry pi to collect the information of different parameters from the surrounding. DHT11 sensor is used to monitor temperature and humidity of the surrounding. Light Dependent Resistor (LDR) and capacitor arrangement is used for sensing the light condition.HC-SR501 is a passive infrared sensor (PIR) used for motion detection around the robot. A low cost Xpro desktop microphone was used for voice command input. Edup N8553 mini wifi adapter is used for internet connection and remote monitoring capability. The raspberry pi is running Raspbian Jessie which is a debian version of linux. The activation of the system and process monitoring is done by connecting the system via LAN connection using Secure Socket Shell(SSH) service. The communication between robot unit and home control unit is wireless using 434 MHz Radio Frequency (RF) Module. RF Transmitter is part of the robot unit. All the major components of the system are written using Python v2.7 programming language. TowerPro MG996R servo motor is used for the jaw control of the robot. Two 1.5 mm red colored led were implemented to represent the eyes of the robot for visual feedback.

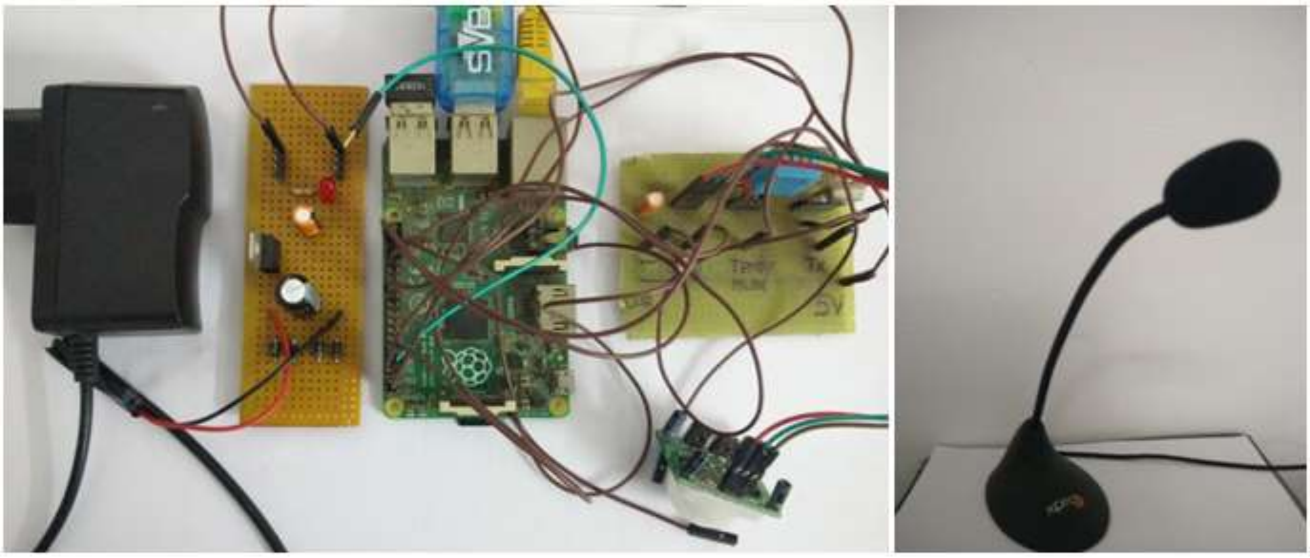


Figure 6 Robot Unit

The second unit of the system is home control unit. In Fig.7 Home control unit is shown which consists of Arduino Uno R3 as the main processing unit. Arduino controls the 8 channel relay unit using ULN2803 Integrated Circuit (IC) chip for controlling different devices. RF Receiver is connected for receiving the input from the robot unit.

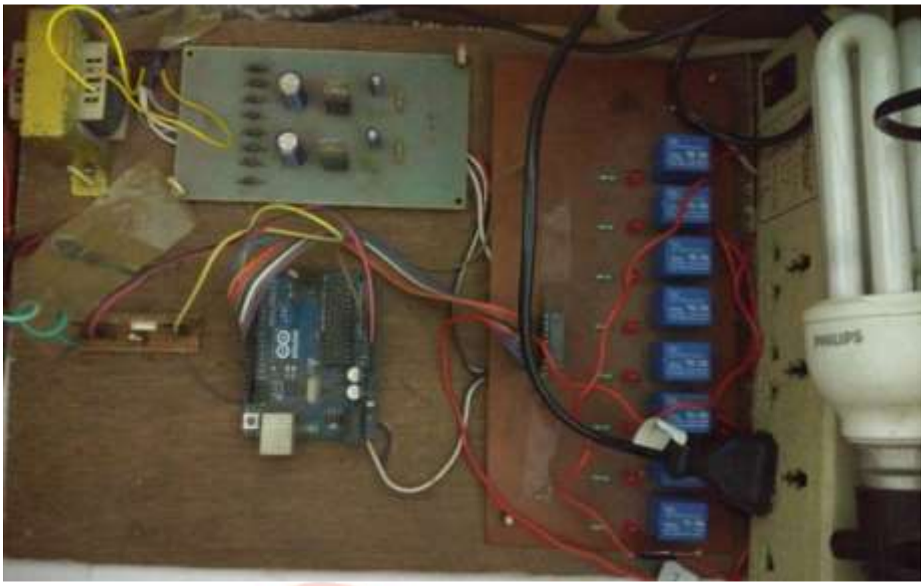


Figure 7 Home Control Unit

Speech Processing

Speech processing is very important part of this project since the input to the robot is through voice commands which are processed in real time. Speech to Text Processing (STT) and Text to Speech (TTS) are both included in the project. Pocketsphinx is an open source continuous speech recognition project by Carnegie Mellon University and is used for converting the voice commands into text data. A corpus of 23 commands is created for the control of robot which included commands related to device control, for example “Turn on the kitchen light” is one such command. Since the robot is context aware it also serves commands like “What is the temperature” and “Is there any nearby movement”. A dictionary was created using the Sphinx knowledge base tool[12] using the corpus. In Fig.8 Snippet from the dictionary file is shown.

```
INTRODUCE        IH N T R AH D UW S
FAN        F AE N
FRIDGE     F R IH JH
KITCHEN    K IH CH AH N
LIGHT      L AY T
```

Figure 8 Snippet from the dictionary file

JSpeech Grammar Format (JSGF) is a platform and vendor independent representation of grammar that can be used by speech recognition systems. Using grammar format is helpful since it restricts the recognition of sentences under a stated scope and helps to increase recognition accuracy of automatic speech recognition system. Snippet of grammar file used in the system is shown in Fig.9.

```
#JSGF V1.0;
grammar commands;

<action> = TURN ON         |
           TURN OFF        ;

<object> = KITCHEN LIGHT   |
           FAN             |
```

Figure 9 Snippet from the grammar file

In order to convert text to speech for the robot to give audio feedback eSpeak open source speech synthesizer is used. Command line and shared library version of eSpeak for windows and linux is available.

V. RESULT AND ANALYSIS

The results and evaluations are presented in this section. Tests were conducted in order to check the performance of the speech recognition system used in the project. Four different users individually recorded all the 23 commands which are part of the corpus. The recordings differed with respect to the speech rate, pronunciation and voice quality. SphinxTrain is an acoustic model trainer library by Carnegie Mellon University used for generating the results [14]. In Table 1 results of four different users interacting with the social robot CASOR is shown. The corpus has a total of 95 words in the 23 sentences as commands. Error is the summation of insertions, deletions and substitutions of words manipulated in the decoding process.

Table 1 Results from the speech recognition system

Users	Total Words	Correct Detection	Insertions	Deletions	Substitutions	Error
User 1	95	83	1	3	9	13
User 2	95	76	2	4	15	21
User 3	95	67	0	15	13	28
User 4	95	81	0	12	2	14

Performance of a speech recognition system can be evaluated using a common metric know as Word Error Rate (WER). WER is calculated by summarizing the total errors in the hypothesis divided by total number of words in the transcription. An error is either an incorrect substitution, deletion or insertion of a word which causes the hypothesis to differ from the original sentence. The WER is derived from the Levenshtein distance, working at the word level instead of the phoneme level [6]. Levenshtein distance can be computed as shown in Equation 1. WER is represented in percentage and can be calculated from the Equation 2.

$$\text{Levenshtein distance} = \text{Substitutions} + \text{Deletions} + \text{Insertions} \qquad (1)$$

$$\text{WER} = \frac{\text{Substitutions} + \text{Deletions} + \text{Insertions}}{\text{Total number of words}} \qquad (2)$$

Accuracy is also a reasonable measure of the decoder’s performance in a speech recognition system [13]. It can be computed using the following Equation 3.

$$\text{Accuracy} = \frac{\text{Total number of words} - \text{Deletions} - \text{Substitutions}}{\text{Total number of words}} \qquad (3)$$

In Table 2 total recording of speech time of every user is given along with the average decoding time. WER and Accuracy values are computed for different users. Accuracy is also measured in percentage.

Table 2 Evaluation of speech recognition system’s performance

Users	Total Recording (in sec)	Average decoding time (in sec)	Word Error Rate (WER)	Accuracy
User 1	72	52.31	13.68 %	86.32 %
User 2	48.38	37	22.11 %	77.89 %
User 3	42.56	29.12	29.47 %	70.53 %
User 4	39.05	27.64	14.74 %	85.26 %

The social robot with the ability of speech recognition and speech synthesis is also context aware. Context awareness is achieved with the help of different sensors that are part of robotic system. Various parameters like temperature, humidity, light condition are sensed every second and the data is represented in real time. Nearby motion is also detected by the robot with the help of PIR sensor.

A python programming based Web Server Gateway Interface (WSGI) micro web framework called Bottle is used to create the backend server to fetch data from all the sensors. The data is refreshed every second and is exposed in the form of Javascript Object Notation (JSON) to the front end. Using web technologies like Hypertext Markup Language (HTML), Cascading Style Sheets (CSS) and Javascript, the front end web interface was designed. Flot js which is jQuery based plotting library was used to generate graphs with continuous stream of data. Fig.10 shows the four different graphs for temperature, humidity, light condition and nearby movement. All these graphs are continuously updated. The different commands related to querying of these parameters are successfully served by the robot. In order to receive most recent data the values of these parameters were constantly updated in a temporary file from which the robot read the sensed values. The temperature measured in degree celsius represents room temperature. The humidity of the surrounding is measured as relative humidity in percentage. Light intensity is measured using an LDR and capacitor arrangement. As the resistance of LDR decreases with increase in incident light the capacitor gets charged fast and indicates bright condition of the surrounding. Dark condition is represented when the capacitor takes more time to charge. PIR sensor used by the robot denotes nearby movement in graph using binary representation. 1 stands for movement detected and 0 shows that no movement is taking place near robot.

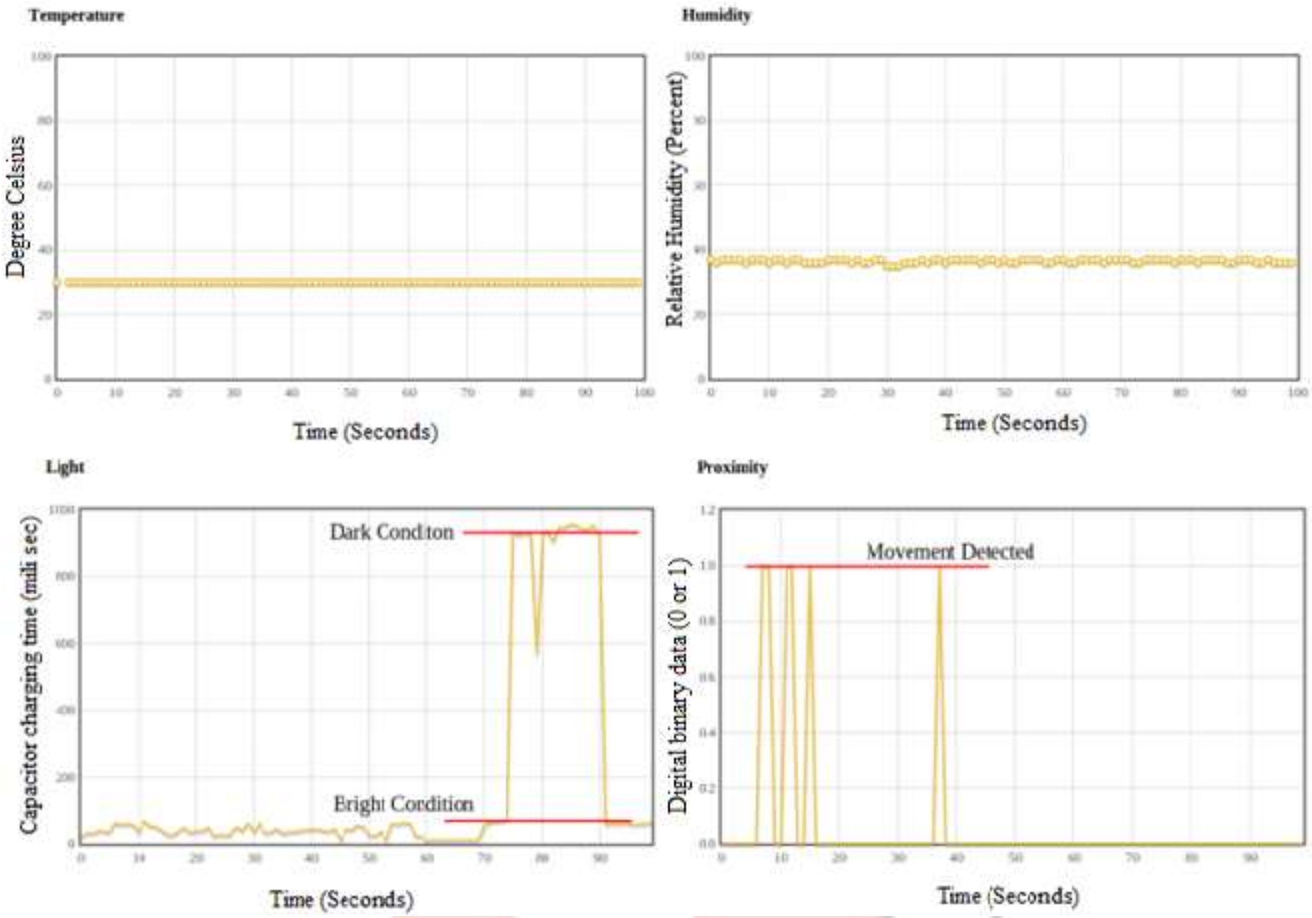


Figure 10 Graphical representations of different parameters in real time

The social robotic system has multiple processes running simultaneously. Fig.11 shows a split linux terminal window accessed via SSH. On the left side of the terminal, continuous speech recognition system is running and a command “Turn on the kitchen light” is detected by the robot. In parallel, a back end server is running to sense data from the sensors every second shown in right side of the split terminal.

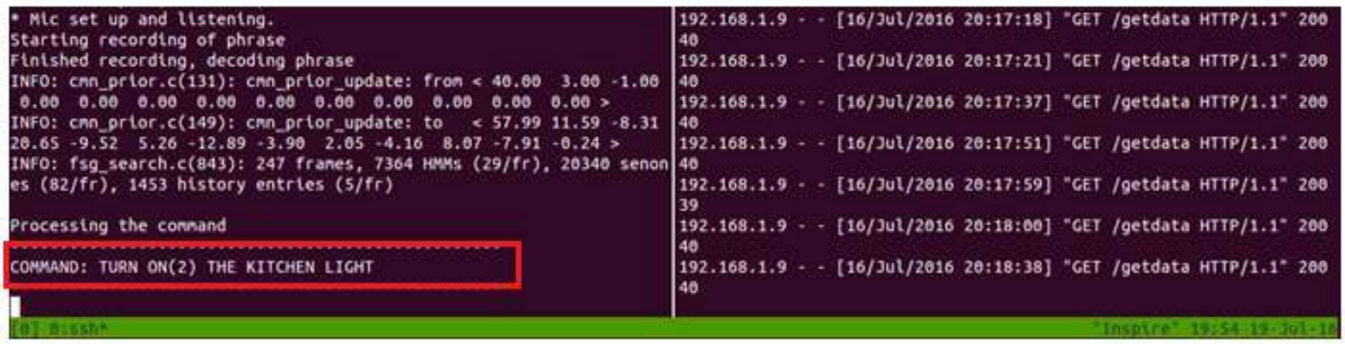


Figure 11 Snapshot of the split linux terminal window using SSH aceess of raspberry pi

After the command is detected by the robotic system, a particular signal is transmitted from the RF transmitter to the home control unit to switch on the specific device. This process can be seen in the Fig.12 where the robot is giving the verbal as well as visual feedback and a device, in this case “Kitchen light” is getting turned on.

The raspberry pi B+ used as the main control unit has limited processing power with 512MB RAM. This factor is visible during speech recognition process and wireless transmission of signals to the home control unit. In spite of multiple processes running at the same time in the robotic control unit, the system performs reasonably well. The speech recognition can be enhanced by adapting the acoustic model and tuning it with the user’s voice. This adaption will overcome the recognition problems which occur due to different pronunciation, speech rate and voice quality. To increase the coverage of the wireless home automation unit, instead of using RF trans-receiver other solutions like zigbee can be used for future enhancement. Since the motive of the project was to create prototype of low cost social robotic platform, off the shelf components were used.

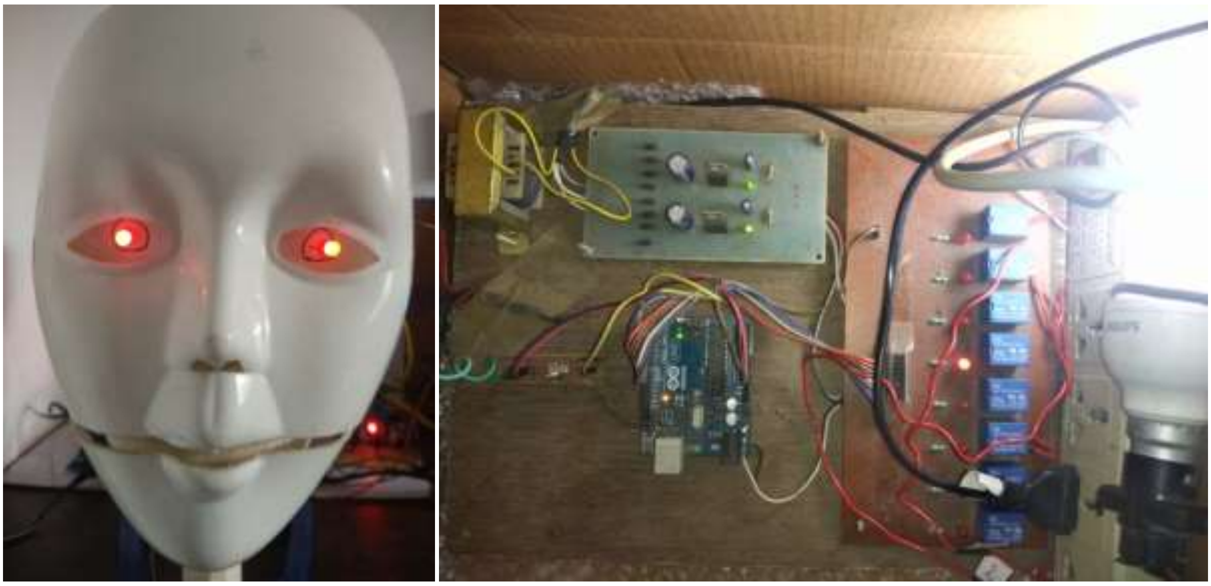


Figure 12 Output from the social robot after decoding the command and turning on the particular device

VI. CONCLUSION

Designing social robots presupposes creating and reinforcing social cues which support social interaction with human beings. To interact socially with a robot means to communicate by referring to social behavioral patterns. The process to build a context-aware social robot for home environment is been discussed. The main focus was to create a low cost robotic platform that can be used to control home devices based on user interaction and sensors input. The various sensor input helped to make the robot intelligent in order to take decisions. Also the system focused to be economical so that many student and researchers can experiment with it.

VII. ACKNOWLEDGMENT

I am using this opportunity to express my gratitude to thank all the people who contributed in some way or the other to the work described in this paper. My deepest thanks to my project guide Professor Rupali Nikhare for giving timely inputs and intellectual freedom of work. I express my regards to the head of Information technology department and the principal of Pillai College of Engineering, New Panvel for extending his support. I am grateful to my college for providing the access of Research lab to work on this project.

REFERENCES

[1] de Graaf, Maartje MA, Somaya Ben Allouch, and Jan AGM van Dijk, "Long-Term Acceptance of Social Robots in Domestic Environments: In-sights from a User's Perspective", Association for the Advancement of Artificial Intelligence(AAAI), 2016.
[2] Ankur Pandey and Rupali Nikhare," A Survey on Recent advance in Social Robotics", International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET), Volume 5, Issue 8, pp.14988-14994, August 2016
[3] Ms Vaishali Wagh, Prof Manisha Wasnik, "Designing of Service Robot for Home Automation", International Journal on Recent Technologies in Mechanical and Electrical Engineering (IJRMEE), Volume 2 Issue 1, ISSN: 2349-7947, PP: 016 – 020, January 2015
[4] Prasanna G. and Ramadass N., "Low Cost Home Automation Using Offline Speech Recognition," International Journal of Signal Processing Systems, Vol. 2, No. 2, pp. 96-101, December 2014.
[5] Michael Ferguson, Nick Webb, and Tomek Strzalkowski, "Nelson: a low-cost social robot for research and education", Proceedings of the 42nd ACM Technical Symposium on Computer Science Education, ACM, pp. 224-229, 2011.
[6] Anusuya, M. A., and Shriniwas K. Katti. "Speech recognition by machine, a review." *arXiv preprint arXiv:1001.2267*, 2010.
[7] Mr. Sangmeshwar S. Kendre, Prof. Dr. Mrs. S. D. Apte, "Voice Controlled Robot using Open Source Hardware Platform Beagleboard", International Journal of Recent Trends in Engineering & Technology, Vol. 4, No. 4, pp.101-105, Nov 2010.
[8] Kerstin Dautenhahn, Chrystopher L. Nehaniv, Michael L. Walters, Ben Robins, Hatice Kose-Bagci, N. Assif Mirza, Mike Blow, "KASPAR-A Minimally Expressive Humanoid Robot for Human–Robot Interaction Research", Applied Bionics and Biomechanics 6.3-4, 369-397, 2009.
[9] Ray, Céline, Francesco Mondada, and Roland Siegwart. "What do people expect from robots?." In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3816-3821. IEEE, 2008.

[10] Huggins-Daines, David, Mohit Kumar, Arthur Chan, Alan W. Black, Mosur Ravishankar, and Alexander I. Rudnicky. "Pocketsphinx: A free, real-time continuous speech recognition system for hand-held devices." In *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 1, pp. I-I. IEEE, 2006.
[11] Feil-Seifer, David, and Maja J. Mataric. "Defining socially assistive robotics." In *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005.*, pp. 465-468. IEEE, 2005.
[12] http://www.speech.cs.cmu.edu/tools/lmtool-new.html
[13] http://cmusphinx.sourceforge.net/wiki/tutorialconcepts
[14] http://cmusphinx.sourceforge.net/wiki/tutorialtuning
[15] http://inmoov.fr/head-and-eyes-assembly-3d-view/

AUTHORS

Ankur Pandey is pursuing M.E in Information Technology with specialization in Artificial Intelligence and Robotics from Pillai College of Engineering, New Panvel. He has 2 years of Industrial experience in Embedded systems and Software Development. His areas of interests are Robotics, Artificial Intelligence, Internet of Things (IOT) and E-commerce.

Rupali Nikhare is working as an Assistant Professor in Computer Engineering Department at Pillai College of Engineering, New Panvel. She obtained M. E. in Computer Engineering and B. E. in EXTC. She has 10 years of teaching experience. Her area of interest is in DSP, IP, CV, Robotics, E-commerce.