ISSN 2063-5346

CONCEPTUALIZATION AND REALISATION OFA HUMANOID SOCIAL ROBOT WITH 5 DEGREES OF FREEDOM AS A COLLEGE GUIDE

Ayana K P[1], Sangeetha Gopan G S[2]

Article History: Received: 01.02.2023 Revised: 07.03.2023 Accepted: 10.04.2023

Abstract

Robotics is currently a rapidly growing branch of technology due to the high demand for automatic and social systems in this modern era. In today's world robots are used to serve humanity not only in industrial automation but also in day-to-day activities. A social robot that can communicate and connect with society would greatly help in solving various needs of humanity which usually a human can do. A socially interactive guiding robot for assisting visitors in the college is designed and presented. The database of questions and answers is created which covers almost all the important locations of a college. When the visitor asks the question about the destination, he/she wants to know, the robot compares it with the set of questions that it is familiar with and responds with a suitable answer. The robot interaction is based on Google API's text-to-speech and speech-to-text conversion. The robot has head motion and jaw motion while speaking with a total of 5 Degrees of Freedom. Face detection is done using LBPH (Local Binary Pattern Histogram) algorithm.

Index Terms—Google API, Reception Assistance, Speech to text and Text to speech, LBPH (Local Binary Pattern Histogram), Face Recognition.

[1]*Robotics And Automation (Dept of ECE) College Of Engineering Trivandrum (CET)* Kerala, India ayanapradeepkumar@gmail.com

[2]*Dept Of Electronics and Communication Engg College of Engineering Trivandrum (CET)*

Kerala, India sangeethagopangs@gmail.com

DOI:10.31838/ecb/2023.12.s1-B.384

I. INTRODUCTION

Social robotics is a rapidly developing paradigm in robotics, where human-robot coexistence happens in such a way as to help humanity in different aspects of life. This includes education, healthcare, daily activities at home and so on. Designing a socially interactive robot is a challenging task. This is mainly because the robot must appropriately interpret the action of humans and respond rightly. Human service robots are developed so that they must be able to smoothly interact with man since they directly offer service to man in most cases. In such a case, the voice recognition and understanding technique is essential. This helps the robot recognize the man's voice and use its specified algorithm to deliver the result to the man in a specific language or multiple languages. This paper presents a reception assistance robot with an interacting system for public service robot application (to be implemented in colleges). The interacting system comprises conversion of speech to text and text to speech and audio speech recognition, for offering smart service using Google API. The robot offers a guiding service to visitors once it detects a person using a camera. Visitors can ask about the route to their destination in college, the robot then matches the place mentioned with the previously provided answer and guides them to their destination.

The robot is specially curated for colleges since its operation is typically based on the structure of a college. A person coming to a college for the first time may find it difficult to locate and reach their desired location. The person may have to ask these locations to random people he finds in the college. These random people may or may not know the reply. So, by this, the visitor may lose his valuable time. This random asking is also not at all possible if he is in a hurry like coming to attend a meeting, for any urgent matter like writing an exam etc. Here comes the application of a guiding robot. One can ask about any location in the college they want to go to, and the robot will reply by giving detailed guidance on the route to the destination. This robot is designed to be placed in front of a college's main block. It is provided with 5 DOF (Degrees of Freedom), including head, jaw and the differential drive motion for the base.

II. RELATED WORKS

Body language is always an essential aspect of communication. This is what should be mimicked for effective human-robot interaction. We humans exchange and convey our thoughts and feelings through our skills like facial expressions, body gestures, voice tone or even by a gaze. Most of the modern robotic platforms currently in use are limited in efficiency to create behaviours of its own, accompany their speech. A neural network based design system can be used which accepts input as audio file and generates upper-

body gestures. The gestures include the head movement, hand movement, and torso movement of the person on humanoid robot called Pepper. This step is done by collecting all required audio features and poses features with help of the data set containing audio and visual recordings that are recorded during participants speaking about a topic they are given. These speaker's personality affects by modifying the movements in the robot [1]. Some other applications of interacting robots include museum guides [4]. The robot voice interaction sys- tem control terminal is provided with a human to computer interaction interface, in which the voice control is achieve by cloud corpus[5]. We can also find well-developed models and systems using RGB and 3D cameras and 20 DOF for robots. Pepper robot has 2 DOF and it does emotion detection based on facial expression [3]. There are several robots namelyTeresa, Double, Pepper, LUCA etc. [6].

The applications of social robots include health and daily care services, teaching assistance, research and application etc[14]. ROS (Robot Operating System) has its own benefits in the areas of vision and conversation. With the use of telep- resence technologies, users can maintain social engagement with other people even while they are in another location by depending on a robotic platform. In order to guarantee realistic interaction from users, many technologies have been applied to socially engaging robots. Robots are used in the field of education, with applications that include vocabulary development, writing aid, and language teaching[6][7]. They can help students learn more and do better academically. In order to increase children's vocabulary development and word- learning skills, the humanoid robot Pepper has been employed in a game environment. Most domestic help robots are still prototypes, and personal assistance robots have yet to be successful. Robots are said to be used in a range of healthcare settings and fall into three categories: social, surgical, and rehabilitative robots[12][19].

III. METHODOLOGY

The methodology adopted for the development of the reception assisting robot shows processes, techniques and procedures for the designing and implementation of the robot.

A. Robot action flow chart

Figure 1 shows the process flowchart of the project. The camera is used for the identification of people approaching the robot. At first, the head moves to the right and checks if any humans are there by using the camera. If there is one who is not familiar with the robot, the body turns towards that position and it turns on the speaker, greets and asks the question "Hello, I am KL10 robot happy to help you. Where do you want to go?". By asking this question, the robot will be ready to service the visitor. So hence now the visitor can ask for the route to the destination to the robot to get guidance. If the destination asked by the visitor matches with the previously set questions by keyword matching, then the corresponding answer will be replied to by the robot as a voice command. But if the question does not match or if the voice is not recognized

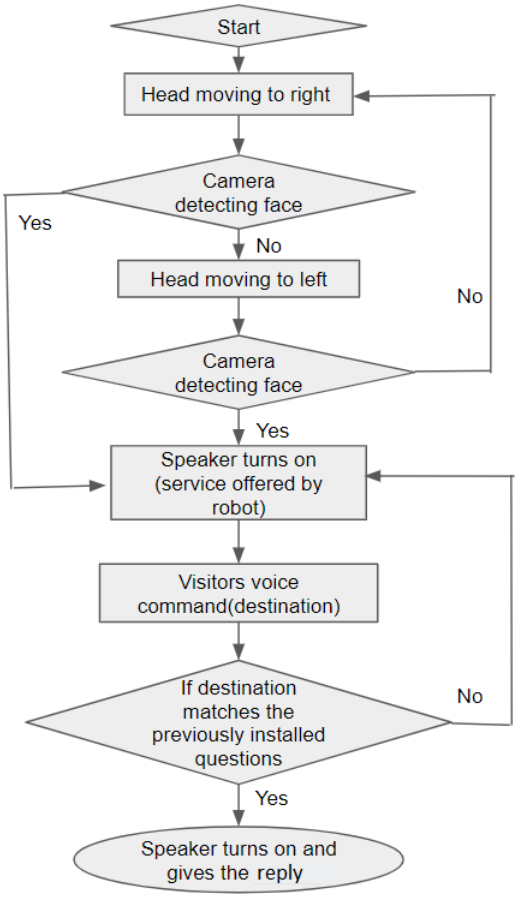

Fig. 1. Diagram of the workflow

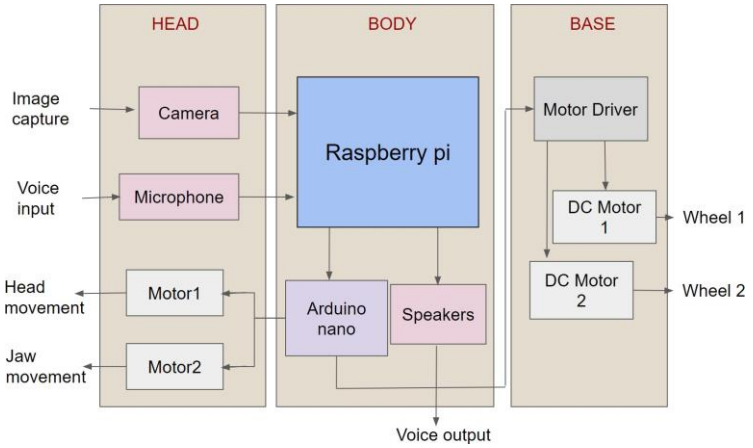

Fig. 2. Block Diagram representation of the entire circuit of robot design

it will again offer the service by asking the question again. If there is nobody on the right-hand side, the head rotates to the left and tries to locate somebody who is new to the place. The remaining procedures are the same. So, the robot will be continuously rotating its head right and left. A block diagram representing the entire circuit can be seen in Figure 2.

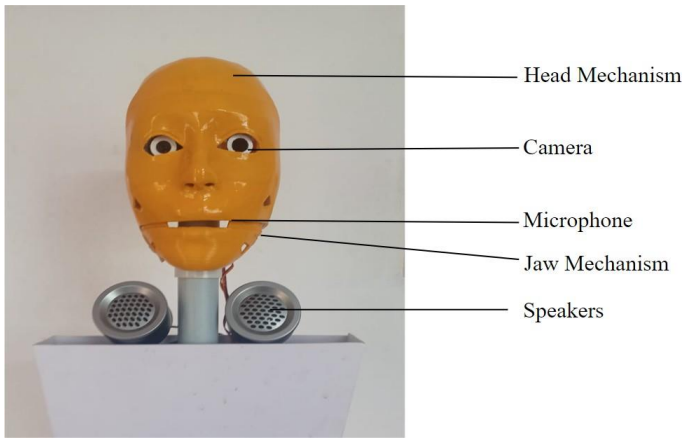

Fig. 3. Robot head hardware setup

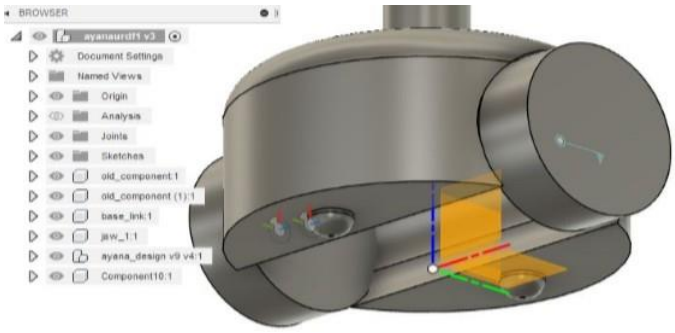

Fig. 4. Robot base hardware setup

IV. Design

A. Head and jaw design and movement

The head of the robot is designed in such a way that it has a rotating joint. Robot's head will be on a continuous left- to-right movement from 90 degrees to 0 degrees and again from 0 degrees to 180 degrees and again back from 180 to 90 degrees. This is just like a surveillance robot trying to detect a face in this whole range. If the robot is detecting a face on the left the body may turn to the left and offer the service. If no visitors are identified on the left it rotates the head and checks the right side, and if it locates a person the robot turns right and offers guidance.

A jaw motion is also provided while the robot interacts with people. A revolute joint is given between the jaw and head so that its mouth would open and close like that of a human while speaking. This is to give aesthetic attraction and to make it more humanoid. Together the head and jaw motions make the head together have 2 DOF. PLA (Polylactic acid) material is used for the 3D printing of the head. The head design adopted here is Inmoov. The robot head hardware setup is shown in Figure 2.

B. Robot base design and movement

The base is designed to turn the robot's body according to the visitor's location. The base is hence given two wheels and two castor wheels are also provided. So, the base has 3 DOF. The base platform provides all the support for the robot's vertical holding and motion. The design is done using Fusion 360 software, it is shown in Figure 3.

V. Algorithm development

A. Questions and answers

It is given a set of most probable questions that a visitor to the college may ask to the robot and corresponding answers are also given. By understanding the voice command from the visitor, the system matches the question with the previously given set of answers and replays the corresponding answer. The robot is designed to answer the queries of the visitor after listening to his/her voice. The Robot interaction code consists of Speech conversion to text, Keyword matching codes and Text conversion to Speech code. The application programming interfaces (APIs) known as Google APIs were created by Google to facilitate communication with and integration of Google Services with other services. These include Google Search, Gmail, Translate, and Google Maps, as examples. These APIs allow third-party programs to utilize or enhance the functionality of the current services.

B. Speech conversion to text

The recorded speech of the visitor is converted to text by using the Google API. Speech is needed to be converted from physical sound form to electrical signal form using microphone. It is then converted to digital data form with an analog to digital converter. Once this is done different models can be used to transcribe this audio to text.

C. Keyword Matching

Two lists were created in the Python program. One is the list of keywords containing questions, and the other is the list of answers to the corresponding question keyword. The text that gets after conversion from the visitor's voice is then matched to the first list and if it is found matched, the corresponding index number is recorded and the answer with the same index number in the second list is printed as the reply answer.

D. Text to speech conversion

Pyttsx3 provides a facility to convert text files they are receiving into different languages such as English, German, French etc. This includes international, national as well as local languages. We can also play the audio speakers in a fast or slow mode as well as in male or female voices of different age groups. These codes will convert the text to speech and store the file in the memory as an MP3 file and play this file after that.

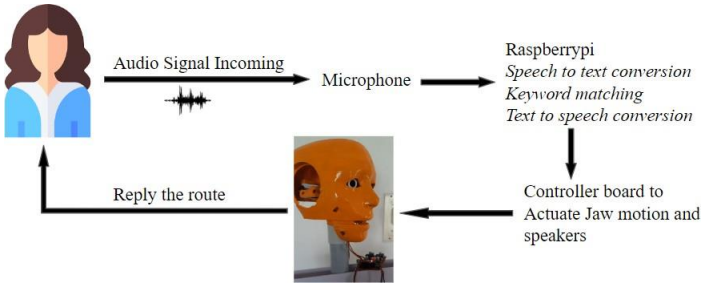

Fig. 5. Schematic representation of robot interaction

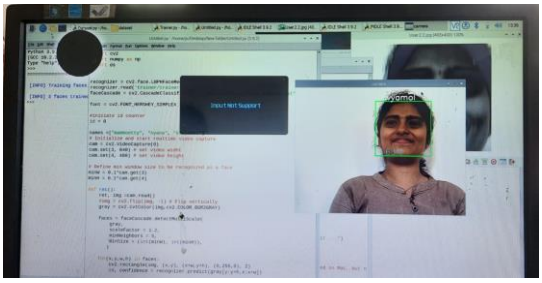

Fig. 6. Face identification demonstration

Fig. 7. Robot interaction demonstration

E. Face detection testing

A 5Mp Web camera is used for image capturing. LBPH (Local Binary Pattern Histogram) is the algorithm adopted to recognize the face. Faces of students are collected live, and about 100 images of a single person are collected and trained using the LBPH training algorithm. LBPH is preferred over deep learning because of its performance. LBPH is faster than deep learning though it needs more data set for being accurate. When implementing this in a college it is possible to use the dataset of students, teachers and staff which is already available in the college database directly. So that the robot can avoid offering service to them and concentrate on visitors or newcomers. Figure 5 shows the working code for the robot identifying the face.

VI. Experimental Results

Result 1: At the first result, it is found that the robot could identify a person who is not in the data set, and it is able to offer the service to the person by considering him as a newcomer to the college.

Result 2: Next a person who is there in the data set has also come near the robot. Now the robot identified that the person is someone who belongs to the college itself, so waited for 5 seconds if he is saying any Keywords for getting guidance. And if there is no response from the person the robot continues to search for the next person. The robotic head turns 10 degrees each and waits for 1 second for face detection (Figure 6).

Result 3: A place from college that is not in the dataset is asked by a person, Robot successfully replied ”Sorry I didn’t get you please repeat”. And again asked the same this time robot finds the destination is not there in his data set he replied ”Sorry! I cannot help you. Please contact the office”.

VII. Conclusion

A humanoid reception assistant social robot with 2 DOF head motions and a total of 5 DOF is designed for interacting with newly coming people to a college. The robot is designed to guide the visitors by telling them the way to reach their specific locations in the college and thus decrease the effort of the people to enquire about routes to unknown people in a new location. The robot’s interaction code is written in Python. The robot’s interaction with the visitors is designed with the help of Google API’s speech-to-text and text-to-speech conversion. The list of probable locations is also created to match with visitors’ desired location and to give a reply. Face recognition using LBPH algorithm is also incorporated to identify visitors from the rest of the people passing by so that assistance will be automatically given to newcomers to the college.

References

[1] J. Ondras, O. Celiktutan, P. Bremner and H. Gunes, ”Audio-Driven Robot Upper-Body Motion Synthesis,” in IEEE Transactions on Cybernetics, vol. 51, no. 11, pp. 5445-5454, Nov. 2021, doi: 10.1109/TCYB.2020.2966730.

[2] K. Youssef, S. Said, T. Beyrouthy and S. Alkork, ”A Social Robot with Conversational Capabilities for Visitor Reception: Design and Framework,” 2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART), Paris / Créteil, France, 2021, pp. 1-4, doi: 10.1109/BioSMART54244.2021.9677852.

[3] B. G. Draghici, A. E. Dobre, M. Misaros and O. P. Stan, ”Development of a Human Service Robot Application Using Pepper Robot as a Museum Guide,” 2022 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), Cluj-Napoca, Romania, 2022, pp. 1-5, doi: 10.1109/AQTR55203.2022.9802037.

[4] Y. Kuno, H. Sekiguchi, T. Tsubota, S. Moriyama, K. Yamazaki and A. Yamazaki, ”Museum Guide Robot with Communicative Head Motion,” ROMAN 2006 - The 15th IEEE International Symposium on Robot and

Human Interactive Communication, Hatfield, UK, 2006, pp. 33-38, doi: 10.1109/ROMAN.2006.314391.

[5] G. Zheng, S. Bi, H. Min, K. Yang and Y. Zhang, "Design and Implementation of Chinese Speech Robot Control System Based on Android Embedded Platform," 2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), Honolulu, HI, USA, 2017, pp. 498-503, doi: 10.1109/CYBER.2017.8446468.

[6] K. Youssef, S. Said, S. Alkork, and T. Beyrouthy, "A Survey on Recent Advances in Social Robotics," Robotics, vol. 11, no. 4, p. 75, Jul. 2022, doi: 10.3390/robotics11040075.

[7] Sarrica, Mauro, Sonia Brondi, and Leopoldina Fortunati. "How many facets does a "social robot" have? A review of scientific and popular definitions online." Information Technology People 33.1 (2020): 1-21.

[8] Duffy, Brian R. "Anthropomorphism and the social robot." Robotics and autonomous systems 42.3-4 (2003): 177-190.

[9] Kuno, Yoshinori, et al. "Museum guide robot with communicative head motion." ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication. IEEE, 2006.

[10] Breazeal, Cynthia. "Toward sociable robots." Robotics and autonomous systems 42.3-4 (2003): 167-175.

[11] Li, Haizhou, Cabibihan John-John, and Yeow Kee Tan. "Towards an effective design of social robots." International Journal of Social Robotics 3.4 (2011): 333-335.

[12] Li, Jamy, and Mark Chignell. "Communication of emotion in social robots through simple head and arm movements." International Journal of Social Robotics 3.2 (2011): 125-142.

[13] Shaw-Garlock, Glenda. "Looking forward to sociable robots." International Journal of Social Robotics 1 (2009): 249-260.

[14] Šabanović, Selma. "Robots in society, society in robots: Mutual shaping of society and technology as a framework for social robot design." International Journal of Social Robotics 2.4 (2010): 439-450.

[15] Korn, Oliver, ed. Social robots: technological, societal and ethical aspects of human-robot interaction. Berlin/Heidelberg, Germany: Springer, 2019. Germany, 2019.

[16] Karar, A. S., S. Said, and T. Beyrouthy. "Pepper humanoid robot as a service robot: a customer approach." 2019 3rd International Conference on Bio-engineering for Smart Technologies (BioSMART). IEEE, 2019.

[17] Bennett, Casey C., and Selma Šabanović. "Deriving minimal features for human-like facial expressions in robotic faces." International Journal of Social Robotics 6.3 (2014): 367-381.

[18] Yoon, Youngwoo, et al. "Robots learn social skills: End-to-end learning of co-speech gesture generation for humanoid robots." 2019 International Conference on Robotics and Automation (ICRA). IEEE, 2019.

[19] Ko, Woo-Ri, and Jong-Hwan Kim. "Behavior selection of social robots using developmental episodic memory-based mechanism of thought." 2018 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia). IEEE, 2018.

[20] D. Rodr´ıguez-Guerra, G. Sorrosal, I. Cabanes and C. Calleja, "Human-Robot Interaction Review: Challenges and Solutions for Modern Industrial Environments," in IEEE Access, vol. 9, pp. 108557-108578, 2021, doi: 10.1109/ACCESS.2021.3099287.

[21] L. Best-Rowden and A. K. Jain, "Longitudinal Study of Automatic Face Recognition," in IEEE Transactions on Pattern Analysis and Ma- chine Intelligence, vol. 40, no. 1, pp. 148-162, 1 Jan. 2018, doi: 10.1109/TPAMI.2017.2652466.

[22] Moreno, Plinio, et al. "Vizzy: A humanoid on wheels for assistive robotics." Robot 2015: Second Iberian Robotics Conference: Advances in Robotics, Volume 1. Springer International Publishing, 2016.