Embelias @Home 2024 Team Description Paper

Piyush Khanna, Shubham Shukla, Atul Kumar, Harsh Chillar, Govind Panwar, and Harsh Jain

KIET Group of Institute, Ghaziabad, Delhi-NCR
https://www.anushkathehumanoid.com/robocup

Abstract. This paper describes the social humanoid robot Anushka of team Embelias for the 2024 RoboCup@Home Challenge.
The research work explores a novel approach of approximating distance using computer vision, engineering process of a pseudo-OS for task handling and prioritizing, the usage of robots to recognize and remember faces in real-time, and their ability to respond to major human languages with versatility, showcasing the potential for culturally inspired AI systems and robotic receptionists.

Keywords: Robotic Receptionist · Distance Approximation using Computer Vision · Speech Synthesis

1 Introduction

Team Embelias is the winner of Humanoid Robotics Championship '23 at DTU (formerly DCE, India) and aims to participate in the 2024 RoboCup@Home being held in Eindhoven, The Netherlands. The team is based at the KIET Group of Institutions, Delhi-NCR and has developed the robot in collaboration with TBI-Ghaziabad.

Anushka is the first robot in the world to be based on Vedic principles of mind's judgement and thinking procedure. Beyond traditional aspects, it integrates IoT for versatile applications, from home automation to remote operations. It is able to recognize and respond to all the Indian languages as well as major foreign languages. It can recognize objects and faces and can remember faces in real-time. It is able to effectively communicate using over fifty types of hand gestures.

This paper describes the overall system architecture, manipulation capabilities, IoT capabilities and novel approach for task prioritization using pseudo-OS, and distance approximation with computer vision without depth sensing cameras.

2 Overall System Architecture

2.1 BOSS: Bot Operations Scheduling System

The proposed system modules need to interact with each other for data transfer and inter-program communication. Also, there needs to be a Master Module that keeps check over all the systems working correctly. It should keep check whether an error has occurred in any of the systems or separate modules and should be intelligent enough to resolve the errors itself or notify the user about it.

Thus, A Master program with a proposed name of "BOSS: Bot Operations and Scheduling System" is designed to act as the master program for the modules such that an ecosystem of performance is built where each program is connected to one another and operates independent of the Operating System they are separately dependent on. The BOSS program also acts as a virtual machine for smooth functioning of all the modules together.

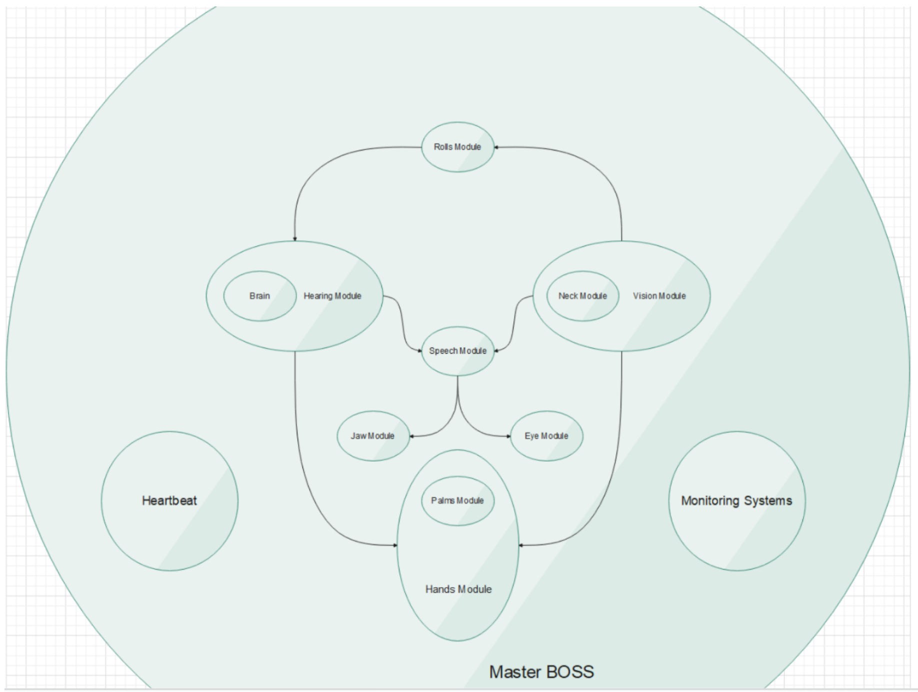

Fig. 1. The BOSS module receives sensing data from all the input modules and smartly decides which output module to command, automatically prioritizing tasks in a background thread.

2.2 Hardware

Anushka uses a HP ProDesk 600 G3 Mini, Intel® Core™ i3-6100T Processor 3M Cache, @3.20 GH, 16 GB DDR-4 RAM as a central processor. Various joints of the robot have 7kgcm, 15kgcm, and 30kgcm servo motors for effective torque. The entire system can be run on a 35C Li-Po battery. The robot has speakers on its back, scapular region, and on the base of moving mechanism. An ANC microphone has been installed on its sternum region.

The skull of the robot has been used by employing the Inmoov skull and the moving mechanism has been made by modifying the NexRobotcis Firebird-XII by removing its on-board computer, and controlling its motors using a cytron driver.

Fig. 2. Anushka, the robot without its clothing, performing a hand-shake on detecting a stranger.

2.3 Communication and Task Flow

Anushka follows a sequential process for judging the task and segregating a command from normal conversation. It can then take suitable actions to complete tasks ranging from Weather prediction, Home Automation to Following a Human using computer vision or automatically making its way, avoiding obstacles.

3 Perception Capabilities

3.1 Distance Approximation using single 2D camera

The robot uses a novel approach for distance approximation using the

Fig. 3. Anushka, the robot saluting and welcoming the Cheif Minister of Utter Preadesh for his visit to KIET Group of Institutions.

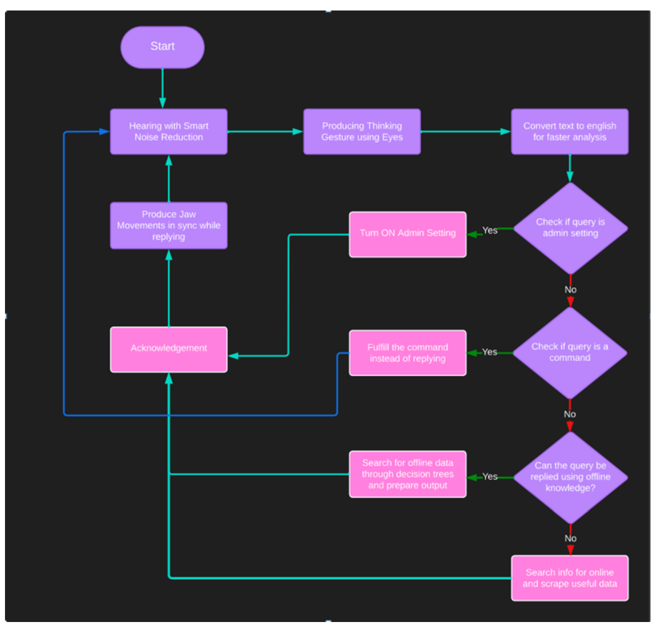

Fig. 4. The robot goes through a segregation process to determine its task.

Len's Formula .

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \tag{1}$$

By using an object whose actual width is already known (u) and the size of the image on the screen (v), the focal length of the camera (f) can be calculated using the lens formula.

This formula can thus be used to determine the distance between the camera and the actual object by placing the values of (f) and (v), the size of the object on the computer screen.

This feature is implemented in the Anushka Robot to follow the tagged person as well as maintain a formal distance from the person she follows.

3.2 Speech Recognition

The robot uses open-source python libraries for online speech recognition such as speech_recognition with recognize_google and googletrans. Offline speech recognition systems such as snowboy and vosk were tried but rejected due to their limited usage and increased local system disk space for storing the data sets and language models. Further, local processing was avoided for faster operation as well as higher accuracy.

Anushka employs NLP algorithms like Sentiment Analysis, Stemming and RAKE to process and understand human language. These algorithms involved tasks like tokenization, part-of-speech tagging, named entity recognition, and sentiment analysis. By utilizing NLP, Anushka could interpret and extract meaning from spoken or written text, enabling it to understand user commands, generate appropriate responses, and engage in natural language conversations.

A novel approach for 'question cache-ing' allows the robot to store the questions and their answers locally based on their frequency of being asked in normal conversations.

3.3 Face Recognition

Novel approach for identifying and prioritizing humans in real-time has been applied to the face recognition system to mimic human behavior. The face recognition system uses deep learning Convolutional Neural Network for first creating face embeddings and storing them in local database.

This database is routinely checked for repeating faces and their mapped names. Priority is given to people based on their number of times meeting the robot and time since last meeting. On a routine-base dumping, faces are "forgotten" if their priority is less, that is, the robot has met the person only once/twice and long time ago.

Face recognition using computer vision typically involves several steps. Here's a high-level overview of the process:

Face Detection : The first step is to detect and locate faces within an image or video frame. Various face detection algorithms, such as Haar cascades, Viola-Jones algorithm, or deep learning-based methods like convolutional neural networks (CNNs), are used to identify potential face regions.

Face Alignment : Once the face is detected, it undergoes alignment to normalize its position and orientation. This step helps in reducing variations due to head pose, scale, and rotation, making subsequent analysis more robust. Landmark detection and geometric transformations are used for face alignment.

Feature Extraction : The next step is to extract distinctive features from the aligned face. Features such as local binary patterns (LBPs) and scale-invariant feature transform (SIFT) are extracted using deep learning-based approach of FaceNet. These features capture unique characteristics of the face that can be used for identification or verification.

Face Representation : The extracted features are used to create a compact representation of the face. This representation serves as a numerical summary of the face's features, allowing efficient comparison and matching. Here, the features are encoded into vectors.

Classification or Matching : In this step, the face representation is compared with a database or a known set of faces. In the robot, we use a hybrid of support vector machines (SVM) and k-nearest neighbors (KNN) for classification of matching tasks. The goal is to identify the person associated with the face or verify if the face matches a particular identity.

Recognition or Verification : Based on the classification or matching results, the system can recognize the person if a match is found or verify if the face belongs to a specific individual.

The final decision is based on a threshold value of the statistical analysis of two determining factors: The last-seen time of that face, and the number of times that face has been seen by the robot.

3.4 Gesture Recognition

The vision module of the robot is able to classify various gestures made by the users along with their detection confidence to suitably determine and remove false positives.

On the confirmation of a gesture being made to the robot, it can respond back by mimicking the same gesture, using its hands module, or by responding to the gesture through speech. The robot employes gesture recognition algorithms, which used techniques like template matching and technologies like

TensorFlow and MediaPipe to identify and interpret hand gestures made by individuals. These were further trained on Teachable Machine to create datasets and classes to be fetched by Classifiers for real-time classification of gestures, allowing Anushka to respond to specific gestures with corresponding actions or gestures of its own.

4 Conclusion

In this paper we have given an overview of the software and hardware to be used by Team Embelias at the 2024 RoboCup@Home challenge. We introduce our swiftly developed robot with customized pseudo-operating system, and the overall architecture designed to tackle the challenges of the competition. Looking forward, we are enthusiastic about the potential impact of our contributions on the landscape of robotic research and development.

References

1. Nourbakhsh, I. R., J. Bobenage, et al. (1999). Affective mobile robot educator with a full-time job, Artificial Intelligence, 114, (1-2), 95-124.
2. Ioannis Iossi_dis1, Christoph Theis, Claudia Grote, Christian Faubel, and Gregor Sch¨oner, "Anthropomorphism as a pervasive design concept for a robotic assistant", In Proceedings Of The IEEE/RSJ International Conference On Intelligent Robots And Systems (IROS 2003), 2003.
3. Scott A. Green, Mark Billinghurst, XiaoQi Chena and J. Geoffrey Chasea, "Human-Robot Collaboration: A Literature Review and Augmented Reality Approach in Design", International Journal of Advanced Robotic Systems, ISSN 1729-8806, pp. 1-18, Vol. 5, No. 1 2008.
4. Cadence L Sidner, Christopher Lee, "An Architecture for Engagement in Collaborative Conversations between a Robot and Humans", Mitsubishi Electric Research Laboratories, Dec-2003.
5. Lingling shi, Hiranya Jayakody, Jayantha Katupitiya, "Coordinate control of a Dual-Arm Space Robot", IEEE Robotics and Automation Magazine, Isuue no.4, Volume 25, pp. 76-85, 2018.
6. Polychronis Kondaxakis, Khurram Gulzar, Stefan Kinaure, Iasonas Kokkinos, Ville Kyrki, "Robot Guesturinng for Anchoring Representation", IEEE Transactions on Robotics, pp. 1-15, 2018.
7. Alessandro Albini, Simone Denei, Giorgio Cannata, "Enabling natural human-robot physical interaction using a robotics skin feedback and a prioritized tasks robot control architecture", 2017 IEEE-RAS 17th International Conference on Humanoid Robotics, pp. 99- 106, 2017.
8. Laura Steffan, Lukas Kaul, Tamim Asfour, "Online stability estimation based on inertial sensor data for human and humanoid fall prevention" , 2017 IEEE-RAS 17th International Conference on Humanoid Robotics, pp. 171-177.
9. Erik Berger, David Vogt, Nooshin Haji-Ghassemi, "Inferring guidance information in cooperative human-robot tasks", 2013 IEEE-RAS 17th International Conference on Humanoid Robotics, pp. 124-129, 2013.
10. Katsu Yamane, Akihiko Murai, "A Cooperative study between Humans and Humanoid Robots", Humanoid Robotics, pp. 1-2-3.

11. Shweta Joshi, Kajal Rathod, Celeste Gudiwada, Mr. Prashant Titare, Dr. D. G. Khairnar, "Motion Based Message Conveyer for Paralytic or Disabled Person", International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653; Volume 8 Issue II Feb 2020.
12. S Behnke, "Humanoid Robots-From Fiction to Reality", Human-Robot Interaction, Issue no. 4, Volume 08, pp. 5-9, 200.

5 Appendix

Name of team: Team Embelias
Members: Shubham Shukla, Atul Kumar, Piyush Khanna, Harsh Chillar, Govind Panwar, Harsh Jain
Contact Information:

– Shubham Shukla: `shubham.shukla@kiet.edu`
– Atul Kumar: `atul.kumar.ece@kiet.edu`
– Piyush Khanna: `piyushkhannavb@gmail.com`

Website: `https://www.anushkathehumanoid.com/robocup`
Hardware:

– Firebird-XII base: On-board computer removed
– Onboard computer: HP ProDesk 600 G3 Mini, Intel® Core™ i3-6100T, Windows Subsystem for Linux
– Active Noise Cancellation Microphone
– Logitech C310 HD Webcam

External Hardware:

– Desktop Workstation
– iBall Bluetooth Speaker

Software:

– OpenCV
– Tensorflow
– google_recognize
– cvzone
– Solidworks
– YOLO