Pak. J. Biotechnol. Vol. 15 (Special Issue) Pp. 1-7 (2018) Manoharan P.A. et al.,
www.pjbt.org, PISSN 1812-1837, EISSN 2312-7791

DESIGN MODELING AND FABRICATION OF HUMAN-HUMANOID ROBOT COMMUNICATION

Manoharan P. A.[1,] John Maria Vivian[2] , Naveen Winson[2] , Solomon Azariah[2],Gnana Kousalya[2]

[1]Research & Innovation Centre, DRDO,Chennai, India. [2]St Joseph's Institute of Technology, Anna University, Chennai, India. E. mails: pamanoharan@gmail.com, johnmariavivian@gmail.com. naveenw1996@gmail.com, soloazar08@gmail.com, hodecestaffaffairs@stjosephstechnology.ac.in

ABSTRACT

Humanoid robotics is an emerging technology, lot of research work is done, its advanced research and development is being continued to incorporate multiple tasks in the robust environment of robotics system. Human Machine Interface (HMI)is a challenging task, which depends upon the critical applications required. Study, design and development of a novel Human face recognition and Speech synthesis essential for the humanoid robot, is carried out in this research project. Voice synthesis and recognition, voice to text conversion is interfaced to the database. Important feature of facial recognition is carried out, using principal component analysis (PCA) algorithm called, Eigen-faces. Prototype model of a humanoid design was simulated, fabricated and tested for the real-time application to interact with the students and faculties. Humanoid Robot (HR) model was validated for its functional capability and enhanced processing speed. HR is capable addressing oral communication of the personnel; HR-HMI, through teleconferencing or video conferencing (VC) at remote terminal.

Index Terms—Humanoid, HMI, VC, PCA, Humanoid Robot, Voice Synthesiser, Supervised Machine Learning, and Man Machine Interface.

I. INTRODUCTION

Humanoid robot is widely designed using 3D modelling software, as the research tools to model and evaluate the perceptual design. Humanoid robots are being developed to perform personal assistance job in industries and home, where they should be able to help out the sick and elderly people to carry out the tasks, which includes impossible tasks of human being. HR is made capable of addressing the users to feel comfortable with their desired solutions by communicating instantly. Humanoid robots are being required to deal with human and machine interactions on numerous complicated stages from physical touch to gestures and spoken languages. Humanoid robots are generally fit into living and factory atmosphere with limited modifications they can be deployed in different environment. From this society should make proper adjustments in order to build humanoid robot communications, which helps make straightforward the robot construction and reduce its cost. Siciliano and Slotine (1991) developed a highly redundant robotic frame work for the multiple tasking and arm tracking model is evolved with differential kinematic equations. Kanoun et al., (2010) presented an approach which is based on kinematic tasks by means of continuous deformation to plan foot placements for a humanoid robot motion and locomotion phase. Sidner et al., (2005) explored the perception that process by an interaction during their perceived connection and the effects of tracking. Rich et al., (2010) implemented an preliminary computational model for recognizing difference among a human and a humanoid robot to exhibit the connection unit involves gesture and speech: directed gaze, mutual facial gaze, conversational adjacency pairs and backchannels. Burgard, et al., (1998) designed a museum guide robot which showed the necessity for the robot's control software to have high speed navigation using first order logic incorporates low-level probabilistic analysis with high-level problem solving embedded system. Ono and Imai (2000) proposed a method for smooth interactions between humans and robots using mindreading mechanism, a representation of expression understanding according to this mechanism. Model of a Mind reading system to Psychological experiments were carried out. Kanda et al., (2002) evolved an autonomous interactive behavior to the robot cognitive science and robotics. HR proto type model is developed, which has enhanced physical expression ability and its related body property. Subsequent article of Kandaet al., (2002) proposed a robot architecture, implemented around 100 behaviors' and 800-episode rules in a humanoid robot. Visualizing tool is used to implement complex system. Petland et al., (Vol. 2185) described about the photo book system tools using three techniques of search base grey scale appearance, 2-Shape and search-based textural properties, proceedings. Zou et al., (2006) extensive study on Sparse principal component analysis (SPCA) using the lasso and proposed an efficient algorithm to understand SPCA for both standard multivariate data and gene expression arrays. Candes et al., (2009) proposed a methodology to make progress both the low-rank and the sparse components exactly by solving a very

suitable convex program called Principal Component Pursuit or data matrix. This methodology detects object in a cluttered background. Mridula et al., (2017) employed Speech recognition technology for the various types of speech and the methods to extract and analyze the structure of the speech and its algorithms for speech-to-text conversion. Preeti and Perneet (2013) reviewed the automatic speech recognition system challenges or issue such as various types of speech classes, speech representation, feature extraction tech-niques, database and performance evaluation. Beck and Teboulle (2009) Presented a new fast iterative shrinkage-threshold algorithm (FISTA) to preserve the computational simplicity of ISTA and experimented its results and found it is theoretically and practically better. De-blurring wavelet-based image demonstrated faster than ISTA by several orders of magnitude. Rajesh Kumar et al., (2011) reviewed automatic speech recognition (ASR) systems using the statistical framework of hidden Markov model (HMM) structural design and working of the standard HMM with its limitations evaluate with advances and refinements of the presented HMM techniques.

HUMANOID ROBOT MODEL

Figure 1. Humanoid Prototype Model

Fig. 1 shows the prototype Humanoid Robot 3D model and its weight to height ratio is 40 kg/ft., fabricated to accommodate all the sensors, actuators and its peripheral devices.

A humanoid robot with its body structure 3D image designed to resemble that of a human is fabricated. Its design encompasses the physical core structure functional capability of human functional and its mobilization like stationary or walking. First phase of this research work is done for the stationary HR to possess six senses and it is limited to the requirements of the HR model. However, the complexity of doing so is deceptively great achievement.

HR prototype model is fabricated using Fibre Reinforced Plastic (FRP) and epoxy to the live human static model, it encompasses the input vision sensors, camera, sensors, GPS, Wi-Fi module, recognizing interaction between human and humanoid robot to exhibit the connection unit involves gesture and speech: directed gaze, mutual facial gaze, conversational adjacency pairs and backchannels. These implanted sensors are interfaced with micro controller unit, which responses human like intelligence to reply with the help of big-data storage system online. HR is capable of communicating in the remote mode operation called Virtual Remote Control (VRC).

II. HR DESIGN AND MODELING TOOLS EMPLOYED**:**.Auto-CAD is used as the 3D modelling tools for the design of HR, few complicated shapes and profile parts were imported ProE. Hardware design tools like Arduino and NEC micro controllers were employed to carry out the multiple tasking simultaneously. Python, C coding and Embedded C programs were used for the algorithm. Supervised machine learning algorithms were also used. This robot controller has a 64-bit microprocessor with 4GB of RAM and a 120GB solid state hard disk which is a non-volatile storage device that stores persistent data on solid-state flash memory. It also has a Arduino16-bit microcontroller to control the lights and various sensors of environmental variables like temperature, pressure.

FABRICATIONAND METHODOLOGY**:** Speech recognition and synthesis tool. Speech recognition is achieved using general-purpose Hidden Mar -kov Models (HMM) speech recognition systems. HMM implies statistical models to deliver a sequence of symbols or quantities, to process as a stationary piecewise/short-time signal. Stationary process of Speech is approximated for a time-scale of 10 mS.

Human voice input is received by HR through mike situated at the HR model's chest. Process of converting modulated voice to analog signal is done at the first stage, noise filtering is done prior to speech recognition in the second stage. Fig. 2 shows HR proto type model with their modules and hardware mounts pictorially given. Mounting

these hardware, were the challenging task and implanted in such away to have modular sections with suitable fasteners to bear the shock and vibration during transition. The speech recognition process of comparing the received signal with the database is emulated to the appropriate text as the output. This text is the main input for the voice interaction program. The converted text which is received is then analyzed and separated into categories for greeting, information, farewell, categorized info is further processed to give an appropriate answer.

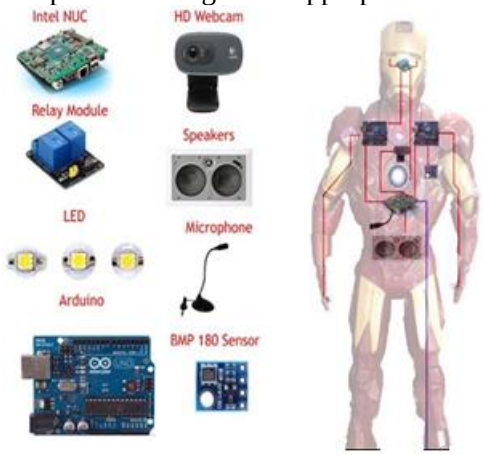

Figure 2. HR Prototype

This system has an automatic learning algorithm which makes it to learn during conversations. The answers that are replied is chosen at random to make it more interactive instead of repetitive, converted back to voice by using a text to speech algorithm or speech synthesis which is the artificial production of human speech. Predefined speech is created by concatenating pieces of recorded module that are stored in a database

A text-to-speech system is consisting of two parts such as front-end and Synthesizer. The front-end has two most important tasks to perform. Initially it converts raw text which contains symbols like numbers and abbreviations into the corresponding written-out words. This process is known as text normalization. The front-end module is used to allocate phonetic transcriptions to each word, and then divides the text into units of phrases, clauses, and sentences. The symbolic linguistic representation is the output produced at front-end module by using phonetic transcriptions and unit information. The synthesizer then converts the symbolic linguistic representation into sound, which includes the computation of the target pitch, rate or speed, which is then imposed on the output speech. The output of the speech synthesizer is an analog signal obtained using another transducer that converts the received analog signals to audible sound.

Chanyaswad et. al., (2016) proposed supervised learning subspace projection method, Discriminant Component Analysis (DCA) for privacy-preserving face recognition applications was achieved. 0Gamage and Lasantha (2014) employed an Independent Component Analysis (ICA) as a feature extraction technique using supervised learning algorithm and Support Vector Machines (SVM).

Scherhag et al., (2017) have studied and review of Supervisor algorithm., Biometric systems morphed face attacks was studied by evaluating the techniques of two databases by printing and scanning digitally morphed images for the vulnerability of face recognition systems. Preet Singh et al. (2016) Compared the Machine learning algorithms Rule-based, Logic-based, Instance-based, stochastic techniques were studied for their complexity. Mapping intentional perception of faces and intentional perception of human body movement frame work for the better performance development of human-humanoid interaction model was deduced by Nagaraj et al. (2017)

The HD web cam is fitted near the chest scans the video feed in the form of images, machine learning process is used to recognize faces in the images and to store the faces in the Saved User Face Database. During communication face recognition is done by comparing with samples obtained from the database if a match is found it greets them, else new registration is made in the database. Data base of samples obtained is shown in the Fig. 3.

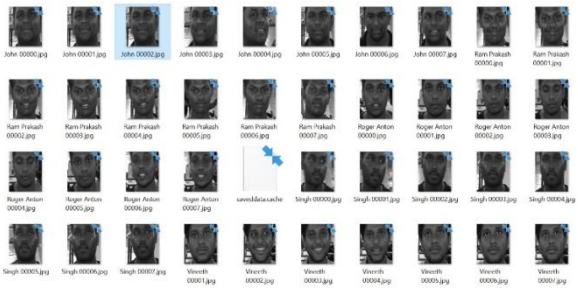

Figure 3: Registered User Face Database

Fig. 3 shows the database consists of multiple images of the same person for improved accuracy. As the quality of the database is very important, from the figure above it is clear that all the images are different from each other. The images of john 0 – john 7 are different from each other, the designed system stores all this data in an organized manner and a cache file is created as shown in saveddata. cache for easy access and faster processing. Face extracted from the video feed and compared the images stored in the data base.

III. HUMAN -HR COMMUNICATION

Human-HR test conditions and Standard operating procedures were arrived, 1. Text and graphic

files are divided, till the text get formatted and styled, 2. Don't use hard tabs, and avoid hard proceeds to only one return at the end of a part, 3. Don't add any sort of pagination wherever in the paper and 4. Don't number text heads-the template.

Table 1. H-HR Communication

S.NO	Number of images in database per class	Number of images compared at a time	Recognition Rate
1.	8	2	95.11%
2.	6	2	91.05%
3.	4	2	86.42%
4.	2	2	76.01%

Table 1 shows the accuracy rate of the Eigen faces algorithm for different number of training images as shown above as the accuracy is directly proportional to the number of training images. Eight number of images has the highest recognizing rate of 95.11% and considered as an efficient database storage.

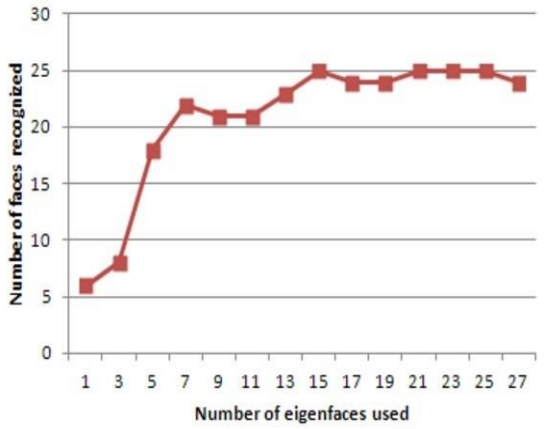

Figure 4: Eigen Face Accuracy

Fig. 4 shows the rate of recognition for different number of eigen-faces. Seven eigen-faces are used to recognize 22 personnel and it continues for the eigen-faces till 15 eigen-faces to 25 personnel face recognition, beyond which the increase in eigen-faces shows a negligible increase in no of faces recognized.

HR also has the ability to differentiate between humans by using the face recognition feature. Facial recognition is achieved by using Python3.5 programming language with Open Source Computer Vision Library (OSCVL). OSCVL and machine learning software library of programming functions mainly aimed at real- time computer vision. This system uses an orthogonal transformation to convert correlated variable into linearly uncorrelated variables algorithm called Principle Component Analysis (PCA). It processes and checks multiple different principal components which is less than or equal to the number of given original variables. This transformation is distinct by the largest possible variance in the first principal component, and following succeeding component is orthogonal to the preceding components which in turn have the highest variance.

An uncorrelated orthogonal basis set is known as resulting vectors. PCA is responsive to the relative scaling of the given original variables. The algorithm used is called Eigen faces. The computer vision problem of human face recognition makes use of the set of eigenvectors which is known as Eigen faces. The covariance matrix of the probability distribution method is used in face images to derive eigenvectors over the high-dimensional vector space. The covariance matrix is built by the set of all Eigen faces images. This covariance matrix turns out dimension reduction to represent the original training images by permitting the minor set of sample images. Classification can be attained by comparing faces with sample predefined set of library called Numpy for numerical operations of the matrix comparison used to match the received image with the basis set.

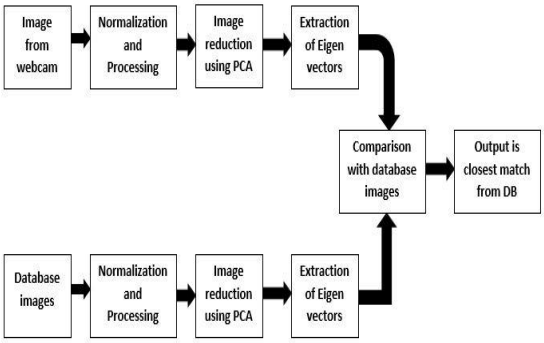

Figure 5: Face recognition

Face recognition process is shown in the block diagram Fig. 5, in which the facial image is normalized, image reduction is obtained using Log Fourier PCA to compare the images from database, error or unrecognized faces are prompted for new entry.

IV. FACE RECOGNITION EFFICIENCY

From the table 1. Accuracy of recognition (a) is proportional to the number of images considered (N).

$$a \ \alpha \ N \qquad (1)$$

As the number of images considered increases the time taken for Recognition (t) also increases.

$$N \ \alpha \ t \qquad (2)$$

Thus

$$a \ \alpha \ t \quad (3)$$

The efficiency of recognition (η) is the number of accurate recognitions (n) in the specified amount of time (T).

$$\eta \ \alpha \ \frac{n}{T} \quad (4)$$

Since

$$n \ \alpha \ \frac{1}{t} \quad (5)$$

Thus

$$n \ \alpha \ \frac{1}{a} \quad (6)$$

Therefore, the efficiency and the efficiency are inversely proportional; hence the optimal efficiency is a trade-off between accuracy and time. Where the Parameters have to be selected based on the application requirements. In the Humanoid robot accuracy is given preference over time as the face recognition feature is used for security purposes.

SPEECH RECOGNITION

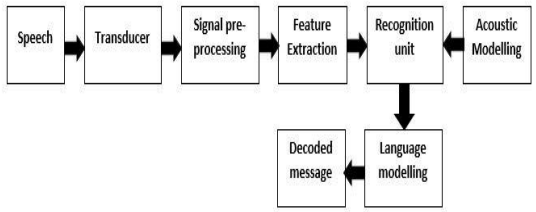

Figure 6: Speech Recognition

Fig. 6 shows the Speech recognition block diagram, in which the mike inbuilt with the HR converts in to text form using speech signal processor, features extraction unit and acoustic modelling process algorithm.

Figure 7: Speech synthesis

Fig. 7 shows the speech synthesis block diagram of HR, in which the text input is utterance composed of phonemes for the appropriate wave form generation input to the transducer to obtain the equivalent speech signal.

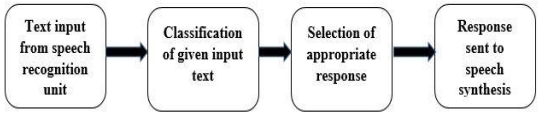

Figure 8: HR Reply Section

HR's reply section block diagram is shown in Fig. 8, during this process communicator input text message is classified and its appropriate responses of speech is sent to the speech synthesis block to answer the communicator.

Fig. 9 shows the HR's response time it was recorded the face recognition time of 2 seconds, communicator response time 1 second and the time between two listening loops is 5 seconds

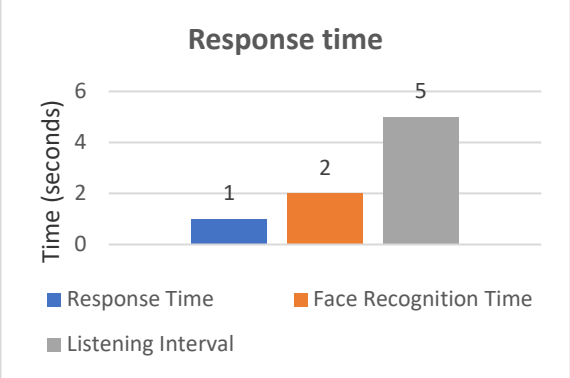

Figure 9: Face Recognition Response Time

V. HUMANOID ROBOT VALIDATION: The humanoid robot has a few special features, it has a database that stores the details of the faculty and students. It can guide the faculty and staff to their respective rooms based on their time table. It first scans the user and checks the database and identifies the user and verifies weather it is a faculty or student using the facial recognition feature. It then asks the user for further instructions, as all the instructions are given through voice commands. Database contains all the secured data. The level of access is based on the amount of clearance given to each user. Currently the humanoid robot has 5 students, 3 teachers, the principal and an admin registered.

The students cannot access the data of the faculty but the faculty can access the data of the students and the principal can view all the data present in the data base. Only the admin has the clearance to modify the data base and the programming of the humanoid robot. The humanoid robot can also be used for research as it can look up information from journals and eBooks. It can also suggest books available on the topic in the Library. The humanoid Robot can also respond to the general queries of unregistered users, it has a map with the layout of the college campus and can guide the users to their destination. The directions are given with respect to the location of the humanoid robot. The humanoid robot is very easy to maintain and upgrade using Python 3.5.3.

HR VIDEO CONFERENCING MODE: The humanoid robot is being accessed remotely through internet, which allows access to the registered users. This allows for the administrator to monitor the surroundings of the humanoid robot and all the people interacting with it. This feature has many uses as it can help with the security and monitoring of the location if any intruder or unauthorized person is detected it can alert the administrator and sound an alarm.

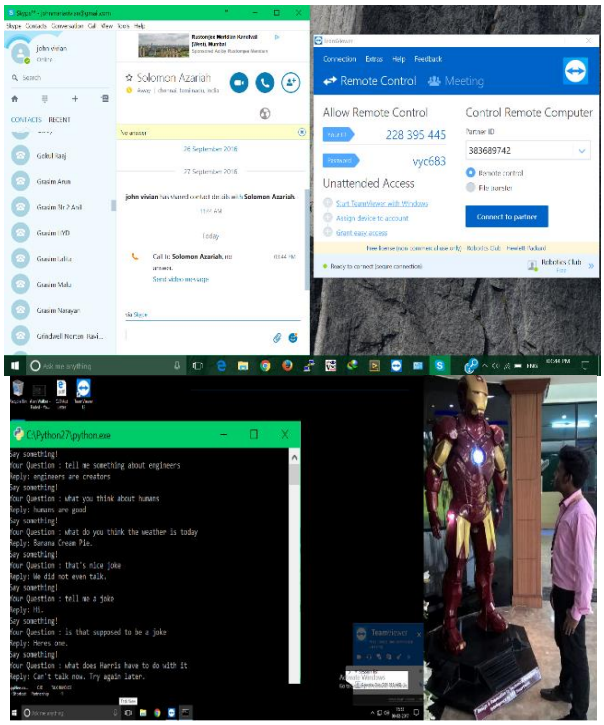

Figure 10: Teleconference Live stream recorded

FACE RECOGNITION

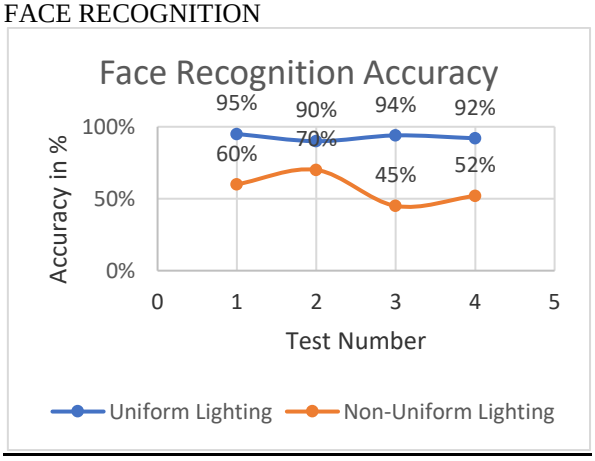

Figure 11: Face Recognition

Figure 11 shows the accuracy of face recognition under different lighting conditions. Test No. 1, shows 60% accuracy of Non-uniform lighting 94% of Uniform lighting and Test No. 3 shows 48% accuracy of Non-uniform lighting and 95% of Uniform lighting. Uniform lighting was obtained by considering HR's chest mounted lamp (350 candela), which shows the highest face recognition. Non-uniform lighting conditions were obtained using three independent sources of lighting each 31 candela.

Figure 12 shows the accuracy of the speech recognition system. The speech recognition accuracy depends on the quality of the microphone used, pronunciation of the user and the algorithm used for recognition.

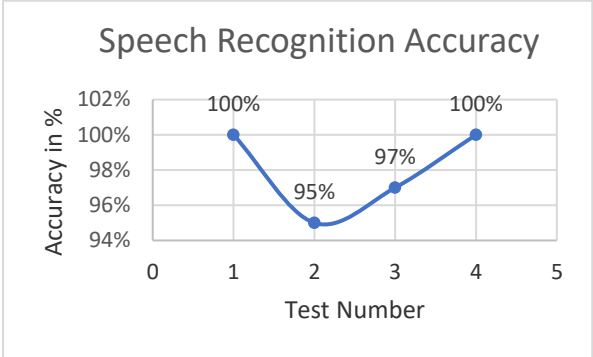

Figure 12: Speech Recognition

The HR uses a high-quality microphone and a complex recognition algorithm. The accuracy of prediction is shown in the graph each test had a set of 20 words which were to be recognized by the HR.

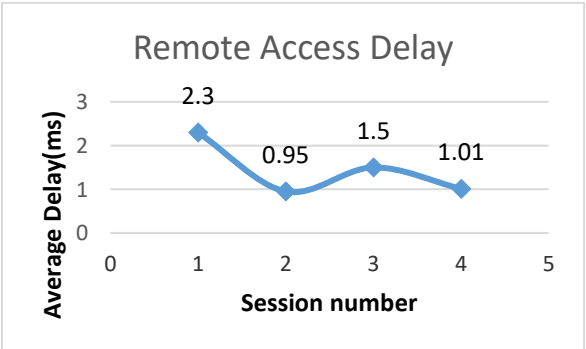

Figure 13: Remote access delay

Remote access delay is the delay that occurs during the operating of the HR remotely this delay is basically the network delay that occurs due to the transmission and reception of data and instructions. Fig. 13 shows the average delay recorded during four different sessions.

VI. RESULTS AND DISCUSSIONS: The humanoid Robot prototype model works well without any errors, replies to any information that is required appropriately. Humanoid Robot is able to recognize and react with the personnel with good conversation. Additionally, HR is capable of incurrect queries to get it confirmed for their information needed by the communicator. The Facial recognition feature works with the highest accu-

racy of 95%. The remaining 5% of the error is due to variations in the light between the images in the database and the given input image. This can be improved by varying the threshold and also by adding more samples with different lighting. The developed Prototype is a stationary HR model is developed for the dynamic bi-pedal robot to make it live interaction in the college campus.

VII. CONCLUSION

This model is used to place as the stationary standing robot to interact with the students, faculties and visitors to guide for their participation in the class rooms, to know about the library books available for lending, general knowledge, to have emergency contact information, to get the security information for the image captured during operation. It has the battery backup to function for 8-hrs during power failure. Future work is to have reduced weight and height of this HR with bipedal movement to mobilize for the real-time environment. Humanoid robots, especially with artificial intelligence algorithms, can be made useful for futuristic design of human assistive, space exploration and other defense applications in the border area to support in the front end of the warfare.

REFERENCES:

Beck A. and M. Teboulle, A fast-iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences 2(1):183–202, (2009).

Burgard W., Cremers A.B., Fox D., Hähnel D., Lakemeyer G., Schulz D., Steiner W. and S. Thrun, The Interactive Museum Tour Guide Robot, Proc. of National Conference on Artificial Intelligence (1998)

Chanyaswad T., J. Morris Chang, M. Pra-teek and S. Y. Kung, Discriminant-component eigen faces for privacy-preserving face recognition, Machine Learning for Signal Processing (MLSP), IEEE 26th International Workshop on 13-16 Sept. (2016).

Emmanuel J.C., Xiaodong Li, Yi Ma and John Wright, Robust Principal Component Analysis, stat web Stanford (2009).

Gamage C. and S.L. Lasantha, Development of a learning algorithm for facial recognition under varying illumination, Information and Automation for Sustainability (ICIAfS), 7th International Conference on 22-24 Dec. (2014).

Kanda T., Ishiguro H., Ono T., Imai M. and R. Nakatsu, Development and Evaluation of an Interactive Humanoid Robot Robovie, IEEE Int. Conf. on Robotics and Automation Pp.1848-1855 (2002).

Kanda T., Ishiguro H.,Imai M., Ono T. and K. Mase, A constructive approach for developing interactive humanoid robots, IEEE/RSJ Int. Conf. on Intelligent Robots and Systems Pp. 1265-1270 (2002).

Kanoun O., Laumond J.P. and E. Yoshida, Planning Foot Placements for a Humanoid Robot: A Problem of Inverse Kinematics. The International Journal of Robotics Research (2010).

Nagaraj B., Reshmi S., Arunkumar R. and M.S. Pradeepraj, A survey on border alert systems for fishermen. Pak. J. Biotechnol. 14 (4) 829-831 (2017)

Ono T. and M. Imai, Reading a Robot's Mind: A Model of Utterance Understanding based on the Theory of Mind Mechanism. Proc. of 7th National Conf. on Artificial Intelligence Pp. 142-148 (2000).

Petland A., Picard R.W. and S. Sclaroff, Described about the photo book system tools using three techniques of search base grey scale appearance, 2-Shape and search based textural properties, proceedings. spiedigitallibrary.org, SPIE Vol 2185, Pp. 35-45.

Preeti S. and K. Parneet, Automatic Speech Recognition. International Journal of Engineering Trends and Technology 4(2): (2013).

Rajesh Kumar A, and M. Dave, Acoustic modeling problem for automatic speech recognition system advances and refinements Par-II. International Journal of Speech Technology Pp. 309–320 (2011).

Rich C., Ponsleur B., Holroyd A. and C.L. Sidner, Recognizing engagement in human-robot interaction, Proceeding of the 5th ACM/IEEE International Conference on Human-robot Interaction Pp. 375-382 (2010).

Subhakala, S., S. Muthulakshmi, A. Geetha, K. Dhanya, Meenakshi Sundara and T.M. Nath, design of smart village using internet of things and cloud computing. Pak. J. Biotechnol. 14(3): 511-513 (2017)

Scherhag U., Raghavendra R., Raja K.B., Gomez-Barrero M., Rathgeb C. and C. Busch, On the vulnerability of face recognition systems towards morphed face attacks, Biometrics and Forensics (IWBF), 5th International Workshop on 4-5 April (2017).

Siciliano B. and J. Slotine, A general frame- work for managing multiple tasks in highly redundant robotic systems, Advanced Robotics 91 and Robots in Unstructured Environments 91-ICAR. 5th International Conference Pp. 1211-1216 (1991).

Sidner, C. Lee, C.D. Kidd, N. Lesh and C. Rich, Explorations in engagement for humans and robots. Artificial Intelligence 166: (2005).

Singh A., Thakur N. and A. Sharma, A review of supervised machine learning algorithms, Computing for Sustainable Global Development (INDIA. Com), 3rd International Conference on 16-18 March (2016).

Zou H., Hastie T. and R. Tibshirani, Sparse Principal Component Analysis. Journal of Computational and Graphical Statistics 15(2): 265-286 (2017).