Feature Based Face Detection in HRI

Nibedita Mishra, Bunil Kumar Balabantaray, Omprakash Sahu, and Bibhuti Bhusan Biswal
Department of Industrial Design, NIT Rourkela, Odisha, India
Email: {nibeditamishra2005, bunil.balabantaray, omprakashsahu}@gmail.com, bbbiswal@nitrkl.ac.in

***Abstract*—Human being is becoming closer and more interested in the collaboration with robots in their daily life. For a better integration with robots, Face detection is a very essential aspect of HRI. In this work, a method for face detection using LabVIEW is proposed. ELM technique is used to filter real face images along with the corner filter. The work is focused on detection of face based on analysis of feature extracted.**

***Index Terms*—ELM technique, feature analysis, face detection, HRI**

I. Introduction

The importance of HRI is rapidly growing due to the increasing use of robots in our daily lives. Robots are also increasingly developed for the real world applications like rehabilitation robots, assistive robots, service robots, educational robots, museum robots, entertainment robots. In HRI, face detection is a very important aspect for developing robots to communicate in a natural way. Face detection is very challenging because of the variation in brightness, inconsistency in measure, position, alignment and posture.

Human being is able to detect and recognize faces in a situation without any struggle or by means of fewer struggles. But, building a robot for human robot interaction having the same ability as human is very difficult as the human face is a very complex pattern. Difficulties like detection of face from any shape in the image, recognition of the face, analysis of facial expressions, and classification based on physical features of the face arise.

In the beginning of 1970s, simple heuristic and anthropometric techniques were used for face detection. But these techniques were inflexible due to the assumptions like background should be plane and the face should be in frontal position. In 1990s, face recognition and video coding in real-time environment started to develop. In 2001, Paul Viola and Michael Jones proposed a real-time object detection framework which is primarily used for face detection. Face detection includes face segmentation, feature extraction and face verification from the undefined and unstructured environment or background.

Manuscript received November 10, 2015; revised August 22, 2016.

II. Related Work

A feature based algorithm for face detection was suggested by Kin Choong Yow *et al.* [1] for different scale, orientation and viewpoint using spatial filtering technique. This algorithm is able to detect the image feature points with the help of spatial filters and also groups them to form the applicant face applying constraints such as geometric and gray level. Then, a probabilistic framework is used to support probabilities and to examine the probability of the applicant as a face. A critical study of the algorithms related with face detection was performed by E. Hjelmas *et al.* [2] along with their possible application areas. The algorithms are classified as feature-based or image-based and are deliberated according the technical approach and performance. In this paper it is cited that the feature-based method is used for face detection in color images in video and image-based is used for grey scale images in static. Y. Pan *et al.* [3] presented an intelligent vision system for face detection by using Ada boost-based Haar-Cascade classifier. The system can detect a human face from any environment and filter out all the fake or non-real face-like images by using Extreme Learning Machine (ELM). Both ELM and AdaBoost techniques were used for improving the accuracy in face detection in the actual real-time speed. Z. Wang *et al.* [4] proposed a human detection system by using a leg detector based 2D laser scanner and a vision based body detector. A probability grid, based on the polar split of the space, was used to merge the data of the two sensors.

An algorithm was suggested by Y. Li *et al.* [5] to detect face that implements skin model and Adaboost face detection algorithm applying Haar feature for face detection in an Instant Message Robot. Matlab program was used for the simulation and optimization of the suggested algorithm. The combinations of Gaussian model implementing K model was used to signify the feature of each pixel in the image. Both single and multiple face detection was done using skin feature and Haar feature based AdaBoost face detection algorithm which has good accuracy and robustness. A behavior System was made by K. S. Ong *et al.* [6] for study of human features and their classification from sensor fusion system. Covariance Intersection (CI) algorithm was used for the fusion of the data obtained from the sensors. They have also suggested a behavior-response system comprising of two fragments (human behavior's inference, robot reaction). They introduced Interaction space, a new spatial zone, which is the combination of

doi: 10.18178/joig.4.2.84-88

personal space and social space. A robot reaction decision algorithm was developed and by them and the evaluation was done for the cognition interaction of the robot. The technique was used by the robot for the person recognition along with proper reaction. Y. Zhang *et al.* [7] recommended an accelerated AdaBoost face detection algorithm and an incremental PCA-based face recognition algorithm for HRI. An image resizing technique and skin color filter combined with a histogram based method was used for face detection in the AdaBoost algorithm. They also explained two vision based algorithms for HRI: the first one was the face detection and real-time tracking algorithm having a less price based on an accelerated AdaBoost classifier and the other one was a face recognition algorithm based on a unique adaptive approach to gradually improve the performance of the face recognition. B. K. Balabantaray *et al.* [8] suggested the finest algorithm (Harris Method) by examining the three major corner detection algorithms (Harris, Wang and Brady Methods) considering accuracy, speed and toughness to noise on Matlab 2012a. They assessed all the algorithms on the images having different types of corners such as L junction, T junction.

A framework in a Mobile Visual Sensor Network (MVSN) for face detection in human-robot interaction was applied by H. M. Do *et al.* [9]. The MVSN was effectively applied for the control of the positioning of the mobile sensors by its head movement. The design was recommended for the surveillance application in real-time environment and was validated for human recognition. B. K. Balabantaray *et al.* [10] proposed an approach, which is a combination of fuzzy partition and wavelet transformation operator, for edge detection in robotic assembly system. They have also used Matlab R2012a for the same purpose. Better result was found by applying the proposed method than the other existing edge detection methods like Canny's, Sobel. A robotic end-effector integrated with multiple sensors was developed by O. P. Sahu *et al.* [11] for the identification of unstructured parts and anonymous environment in industrial robot. They have investigated the state-of-art of the sensor technology for automated assembly. They have used ultrasonic sensor to obtain the dimension, position and weight distribution of the object and vision sensor to know the shape of the surface area. Manish Dixit *et al.* [12] studied the present image recognition methods along with the performance improving hybrid method of the existing feature extraction techniques. They have suggested that for the designing of an image recognition structure considerations should be made on type of image pattern, image feature selection, feature extraction, selection of classifier, and sample selection for training, testing and recognition.

III. METHODOLOGY

Facial Vision System, shown in Fig. 1, is having two segments, i.e., face detection and face recognition. Face detection defines the position and dimensions of human faces in digital images ignoring other things in the environment. Face recognition concludes the uniqueness of the face enclosed in the image area which has been already identified by a face detection system. Face detection and face recognition are challenging tasks due to deviation in brightness, inconsistency in scale, position, alignment (up-right, rotated) and posture (frontal, profile).

Basically, Feature based approach and Image based approach are used for face detection. Low level analysis, Feature Analysis and the Active Shape Models are the steps to be followed for Feature based approach. Low level analysis deals with the separation of visual features using pixel properties. Feature uncertainties are reduced and positions of the face and facial features are determined in the Feature analysis. Active Shape Models is used for the extraction of the complex and non-rigid features.

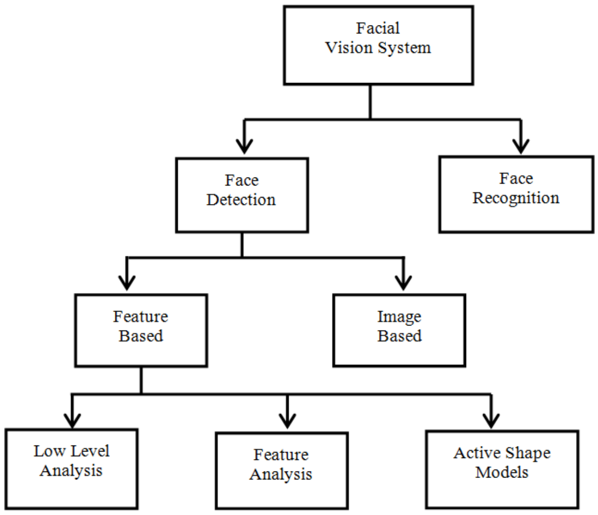

Figure 1. Overview of facial vision system.

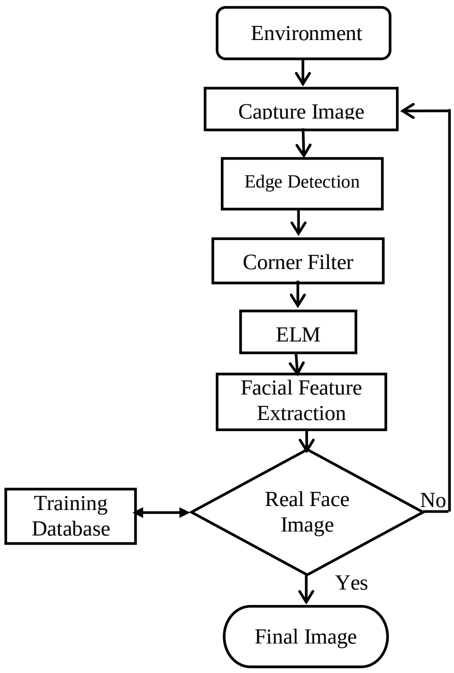

Figure 2. Flow-chart of the proposed methodology

In this work, the face is detected from a real time environment. In order to trace a face, feature based face detection approach is used. Here, only four features of a face are considered for the face detection in a real-time environment. The proposed methodology consists of seven steps and the flow of the process is shown in Fig. 2.

Step 1: Capturing Image: In human robot interaction (HRI), human can interact with robots in many ways like facial expression. Hence, Face detection is very much necessary for communication between human and robot to have an efficient outcome. The image is captured from a real time environment by using a Basler Camera (scA640-70gc). The image is used as the input for the next process for face detection.

Step 2: Edge Detection: The image taken may contain some image of other than human face image present in that environment. Subsequently, by taking the captured image as input, the edge of all the objects is detected during edge detection. The edge of the face is detected by using the hybrid edge detection technique proposed by B. K. Balabantaray, B. B. Biswal [8].

Step 3: Detection of face image: After edge detection, corner filter is used to filter out the non-face components and properties. As the image is captured from the real environment, so many different patterns like notebooks, fan etc. is present in the scene. So it is required to filter only the face like patterns.

Step 4: Detection of real face image: The face like patterns may also have the real face image and the face image from the photograph or calendar hanging in the environment. So, Extreme Learning Machine (ELM) is used to detect the real face images from the image captured and ignores the false or fake face like structures.

The learning or training procedure of the ELM technique used in this work is summarized as below.

$$N = \{ (x^i, t^i) \mid x^i \,!\; R^n, t^i \,!\; R^m, i = 1,...,N\}$$

where, N is the given training dataset

L hidden layers

a_i and b_i are learning parameters of the network

1. Random values to a_i and b_i is assigned for i=1, …, L
2. Hidden layer matrix H is calculated.
3. The hidden-to-output weight is calculated by $W=H^{\dagger}T$, where $H^{\dagger}$ is the Moore-Penrose generalized inverse of H.

Here, ELM is used for the human face detection in the real-time situation or environment. The face like shapes in the captured image is scanned in all the possible scales. For determining the shape as a real or non-real face, ELM algorithm is used here. ELM algorithm is able to classify the faces from different situations like fluctuating lighting condition, different scales of the faces. Detection of real face is a regular and difficult job to prove and test the abilities of ELM. There are also some other objects having the similar characteristics that the real face image is having. Hence, an effective approach to find faces using ELM should be generalized properly for other spatially distinct shape and feature detection problems. The complete procedure for this: each image is considered as a vector in a high-dimensional space, and then ELM is used to the training image datasets to obtain a hidden matrix H. Last the hidden-to-output weights by $W = H^{\dagger}T$, where the elements of T belong to -1 or +1.

Step 5: Facial Feature Extraction: From the different features like face geometry, pixel properties, gray-scale and color properties etc., only the color feature of the image is considered here. The facial features like eyes, nose and mouth are extracted from the real face images detected in the previous step by ELM.

Step 6: Data Acquisition and Database Training: This is a required step for the face recognition from the real time situation as it will be used as the database for an automated system. Hence, in order to recognize the input image, the database should be trained by giving the image as input before the detection of the facial features.

Step 7: Face Recognition and Final Image: The extracted facial features are compared to the face in the database acquired during training. The comparison is made between the color features of the trained image and the real image. After comparison the detected face image is obtained.

IV. Result

In the experimental set up, a Basler Camera (scA640-70gc) and machine vision system (Make: National Instruments) and the LabVIEW is used. The scenario considered in this work is that in an environment containing persons along with other objects. The proposed face detection technique is able to eliminate the other objects and non-real faces. Also, the proposed technique can able to extract and detect the exact person's face if present in the image taken at real time.

The image is captured through the Basler camera and acquired the Image Acquisition function of LabVIEW. Then it is processed by vision assistant and other functions. The features are set for acquired image and processed through a number of functional steps such as edge detector, filters etc.

Hybrid edge detector is used as its detection efficiency is higher than the Canny's edge detector. Then the corner filter is used. Corner detection or interest point detection is a method used in the vision systems for the extraction of different types of features and achieving the focuses of the image. The corners with higher eigenvalue of the image can be found by applying the first-order derivative to each pixel in the image for the comparison of the derivative of its 3×3 pixel region. More corners are found due to complexity in a real face image; the value of all the corners present in the image is compared with a predefined threshold value to consider as a real face image. If the value of the corner present in the taken image is smaller than the value of the predefined threshold, then the image is considered as a non-real face image. However, it is illustrated in various literatures that, generally, various face-like images have many corners which only corner filter cannot completely eliminate from the real face images. Thus, for the improvement in the accuracy of face detection, non-real face images can be eliminated using ELM. ELM is used to estimate the boundary function to separate real face pattern from non-real face pattern.

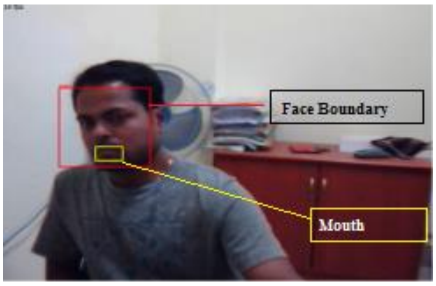

Figure 3. Face boundary and mouth features are matched

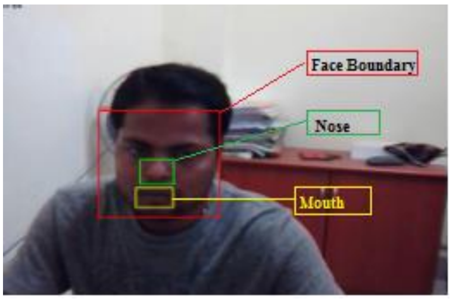

Figure 4. Face boundary, mouth and nose features are matched

The different facial features are compared with the face features stored in the database. Here, the features of face as a whole, eye, mouth and nose is considered. If the features of all four are matched then the face is detected. If only the feature of mouth is matched then the face is not to be considered as the detected face. In the proposed method one by one feature are compared then final detection result is found. As shown in Fig. 3, only mouth and face as a whole is detected, this cannot be considered as a final result. Similarly, face, mouth and nose is detected as shown in Fig. 4, but the features of eyes are not matched. As shown in the Fig. 5, if all patterns are matched, then it is considered to that the real face is detected.

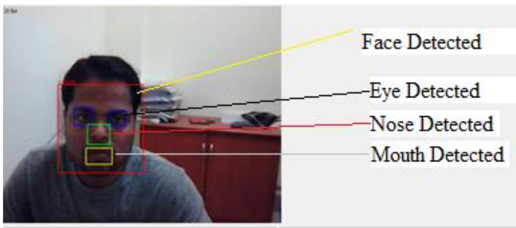

Figure 5. All the four features are matched

V. Conclusion

Basically, there are two different types of communications (verbal and non-verbal) present for the interaction of humans and robots. For facial expression in the non-verbal communication, face detection plays very important role. As the different features like facial expression, lip movement, eye movement etc. is necessary for the communication, we have focused on the feature based face detection.

In this work, a feature based human face detection system is proposed for human robot interaction. Corner filter and ELM technique are used to detect the face from a real environment by considering the color feature. In this work, four elements or features of a face are considered i.e., face as a whole, nose, mouth and eyes. If any one of the facial feature of the captured face image will not match with the information in database, then the captured image will not be considered. Face will be detected if all four elements are matched. The system was tested in real-time environment.

The future work includes face detection considering more facial features along with more than one person. The work also includes the face detection and recognition in a more complicated environment for the improvement in the performance in human robot interaction.

References

[1] K. C. Yow and R. Cipolla, "Feature-based human face detection," *Image and Vision Computing*, vol. 15, pp. 713-735, 1997.

[2] E. Hjelmås and B. K. Low, "Face detection: A survey," *Computer Vision and Image Understanding*, vol. 83, pp. 236-274, 2001.

[3] Y. Pan, S. S. Ge, H. He, and L. Chen, "Real-time face detection for human robot interaction," in *Proc. RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, 2009, pp. 1016-1021.

[4] X. Wang, X. Clady, and C. Granata, "A human detection system for proxemics interaction," in *Proc. 6th ACM/IEEE International Conference on Human-Robot Interaction*, 2011, pp. 285-286.

[5] Y. Li, H. Zhang, and Y. Zhang, "Research on face detection algorithm in instant message robot," in *Proc. International Conference on Uncertainty Reasoning and Knowledge Engineering*, 2011, vol. 1, pp. 220-223.

[6] K. S. Ong, Y. H. Hsu, and L. C. Fu, "Sensor fusion based human detection and tracking system for human-robot interaction," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 4835-4840.

[7] Y. Zhang, K. Hornfeck, and K. Lee, "Adaptive face recognition for low-cost, embedded human-robot interaction," *Intelligent Autonomous Systems*, pp. 863-872, 2013.

[8] B. K. Balabantaray and B. B. Biswal, "Part identification in robotic assembly using vision system," in *Proc. Sixth International Conference on Machine Vision*, 2013, p. 906714-1-5.

[9] H. M. Do, C. Mouser, M. Liu, and W. Sheng, "Human-robot collaboration in a mobile visual sensor network," in *Proc. IEEE International Conference on Robotics and Automation*, 2014, pp. 2203-2208.

[10] B. K. Balabantaray and B. B. Biswal, "Hybrid edge detection technique for part identification in robotic assembly system under vision guidance," in *Proc. 5th International and 26th All India Manufacturing Technology, Design and Research Conference*, 2014, pp. 238-1-238-6.

[11] O. P. Sahu, B. B. Biswal, S. Mukharjee, and P. Jha, "Development of robotic end-effector using sensors for part recognition and grasping," *International Journal of Materials Science and Engineering*, pp. 39-43, 2015.

[12] M. Dixit and S. Silakari, "Intelligent image recognition frame work: An exhaustive survey of face detection & facial feature extraction approaches," *International Journal of Computer Science and Network*, vol. 1, no. 2, pp. 67-73, 2012.

Nibedita Mishra was born in Talcher, Odisha, India, in 1985. She received the B.E. degree in Mechanical Engineering in 2006 and M.Tech. in Production Engineering in 2011 from the BPUT, Odisha, India. Currently she is perusing PhD at National Institute of Technology Rourkela, India.

She has working experience as a faculty from Nov' 2006 to Oct' 2012. She has been published more than 3 technical research papers at National and International levels. Her areas of research include Human Robot Interaction, Assistive Robots and Service Robots.

Ms. Nibedita Mishra is the Life Member of The Indian Society for Technical Education (ISTE: LM 71104). She is awarded with a scholarship from Govt. of India for Ph.D. She has reviewed one paper for Journal of Mechanical Engineering Science.

Bunil Kumar Balabantaray was born in Cuttack, Odisha, India in 1984. He received the B.E. degree in Information Technology in 2005 and M.Tech. in Computer Science and Engineering in 2010 from Biju Patnaik University of Technology (BPUT), Rourkela, Odisha, India. Currently he is perusing PhD at National Institute of Technology Rourkela, India.

He has worked as LECTURER in B.C.E.T. (under BPUT), Odisha, India. Then he worked as assistant professor at DRIEMS (under BPUT), Odisha, India from 2005 to 2010. Then he worked as assistant professor at Synergy Institute of Technology (under BPUT), Odisha, India from 2010 to 2012. He has been published more than 13 technical research papers at National and International levels journal and cof. He published 2 book chapters in the edited book in Springer (1. "Artificial immune system based path planning of mobile robot." In Soft Computing Techniques in Vision Science, pp. 195-207. Springer Berlin Heidelberg, 2012 and 2. Comparison of Edge Detection Algorithm for Part Identification in a Vision Guided Robotic Assembly System." In Soft Computing Techniques in Engineering Applications, pp. 183-206. Springer International Publishing, 2014). His areas of research include mobile robots, industrial robotics, machine vision system, and image processing.

Mr. Balabantaray is the Life Member of The Indian Society for Technical Education (ISTE: LM 50712) and active member of Computer Society of India (CSI), Odisha chapter. He is also a member of SCIence and Engineering Institute (SPIE ID#: 3576832). He is awarded with a scholarship from Govt. of India for doing Ph. D. He has received foreign travel grant from CSIR to attend a conference ICMV 2013.

Om Prakash Sahu was born in Bhilai, Chhattisgarh, India in 1982. He received the B.E. degree in Electronics and Telecommunication in 2005 and M-Tech. in Instrumentation and control system in 2008 from the University of CSVT, Bhilai, India. Currently he is pursuing PhD at National institute of technology Rourkela, India.

He joined as a faculty in RCEE from 2006 to 2012. He has been published more than 18 technical research papers at National and International levels. His areas of research interest include industrial robotics, control system and Instrumentation, Computer integrated manufacturing and automation.

Mr. Om Prakash Sahu is the Life Member of The Indian Society for Technical Education (ISTE: LM 66043), South Asia Institute of Science and Engineering (SAISE: 20150828001), International Association of Engineers (IAENG: 159292), Science and Engineering Institute (SCIEI: 20150827001), and Regular Member Machine Intelligence Research (MIR) Labs. He got the best paper presentation award in BITCON, Durg. He achieved 94.91 percentile in GATE 2006 and eligible for UGC aid. He is awarded with a scholarship from Govt. of India for Ph.D.

Dr. Bibhuti Bhusan Biswal was born in Rourkela, Odisha, India in 1962. He received the B.E. degree in Mechanical Engineering from UCE, Burla, India in 1985. Subsequently he completed his M.Tech, and Ph.D. from Jadavpur University, Kolkata.

He joined as a faculty of Mechanical Engineering at UCE Burla from 1986 and continued till 2004 and then joined National Institute of Technology, Rourkela as Professor and currently he is the Professor in the Department of Industrial Design and also Dean (Faculty Welfare) of NIT Rourkela. He has been actively involved in various research projects and published more than 110 papers at National and International levels including a monogram on Management Information System for ISTE, New Delhi, and 6 Book Chapters and having more than 3 patents. His areas of research interest include Robotics, FMS, CAD/CAM, Mechatronics, CIM, AMT, Virtual Manufacturing, and Production Management.

Dr. Bibhuti Bhusan Biswal is having membership of IE (I: Fellow), ASME (Member), IEEE (Senior Member), ISTE (Life Member), Association for Machines and Mechanisms (AMM: Life Member), ISTAM (Member), International Association of Computer Science and Information Technology (IACSIT: Member) and International Association of Engineers (IAENG: Member). He was the Keynote speaker of ICACC, 2009 (IEEE Conference), Singapore, ICCET, 2009 (IEEE Conference), Singapore, ICROM, 2014 (IACSIT Conference), Nottingham, UK, RDCN-2015, Bhubaneswar, India and many more.