

RGB-D Face Recognition and Gender Recognition from Color Image

Nita Anil¹, Sivadas T Nair²M. Tech Student, Dept. of CSE, VJCET, Vazhakulam, Kerala, India¹Asst. Professor, Dept. of CSE, VJCET, Vazhakulam, Kerala, India²

ABSTRACT: One of the most popular biometric modalities, facial recognition has a very wide range of applications. Different covariates such as illumination, expression and pose exist and need to achieve better recognition rate by overcoming these covariates. In this work RGB-D image is used for face recognition as it contains more information than 2D image thus improves the accuracy. RISE algorithm is used to compute a descriptor from the facial image based on the entropy and saliency features. Geometric facial attributes are extracted from the color image then both the descriptor and attribute match scores are fused for face recognition. In order to increase the security of the system proposed a method to perform gender recognition after performing face recognition. Here investigated Histogram of oriented gradients (HOG) descriptor for gender recognition. SVM classifier is used to recognize the HOG features for gender recognition. The experimental results indicate that the gender recognition using HOG achieves higher accuracy on color images when compared with gender recognition using DWT.

KEYWORDS: Face Recognition; RGB-D; Entropy; Saliency; Gender Recognition; HOG; SVM

I. INTRODUCTION

Face recognition is used for automatically identifying or verifying a person from a digital image or video frame from a video source. A face recognition system identifies a person by comparing selected facial features. Face recognition using 2D images is difficult in the presence of covariates such as illumination, expression and pose. Whereas RGB-D image contains more information thus improves the accuracy of the face recognition. An RGB-D image consists of a RGB image and a depth map (D) and it is shown in Fig.1. RGB image is a 2D color image which provides texture and appearance information. The depth map contains grayscale values which represents the distance of each pixel from the sensor.

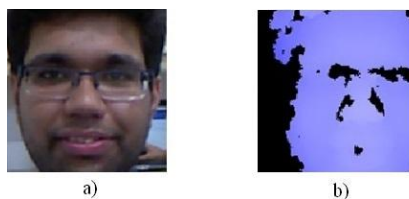


Fig.1. Different modes of capture: (a) RGB image, (b) Depth map

In biometric applications, gender recognition from facial images plays an important role. Recognizing human gender is important since people respond differently based on gender of the person. Further, a successful gender classification approach can be used to boost the performance of many other applications such as person recognition and smart human-computer interfaces.

II. RELATED WORK

In [2] authors used a RGB-D camera to ensure that only the allowed user uses the system. This was the first system to accomplish such objective by using 3D face images. The accuracy of this system is less as it only uses depth information without using the RGB image. In [3] authors shows a method for face recognition across different pose and this is done by obtaining canonical frontal view of face irrespective of the initial pose. In [4] authors make the assumption that for each enrolled subject in the gallery contains only one example face in frontal view. For face recognition across different pose the system generates a set of faces images with different pose angles automatically for

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

each face image in the gallery. In [5] authors address the problem of face recognition under varying expression and used a synthesis based scheme, in which a number of synthetic face images with different expression are created. In [6] authors make use of RGB-D facial image captured from kinect. A descriptor based on entropy and saliency feature obtained from facial image is used for face recognition. In [7] authors use RGB-D-T facial image so by making use of thermal image it gives result invariant to changes in illumination. But it is difficult to interpret thermal images accurately. In [8] authors show how artificially created illumination and rotation helps to cope with illumination and pose variations. In [9] authors show how to recognize face using 3D constrained local model (CLM-Z). In [10] authors presents the first publicly available face database based on the Kinect sensor. In [11] authors show how to estimate the facial landmarks in case of large pose variations.

III. SYSTEM OVERVIEW

Here the Fig.2 shows the working of the entire system. First RGB-D image is preprocessed then RISE algorithm [1] and GA algorithm is applied. RISE computes the texture descriptor from both color image and depth map, Whereas GA computes geometric facial features from color image. RISE algorithm is taken from previous work and the proposed GA algorithm is used. Then in match score fusion combine texture and geometric facial features for Face recognition. Then from the same facial image gender is recognized. Extension to the existing system is done by adding gender recognition after face recognition to increase the security.

For gender recognition the SVM classifier is already trained with male and female HOG features. For testing the gender, HOG feature is extracted from the facial image and then SVM classifier will predict the gender of the person based on the HOG features extracted.

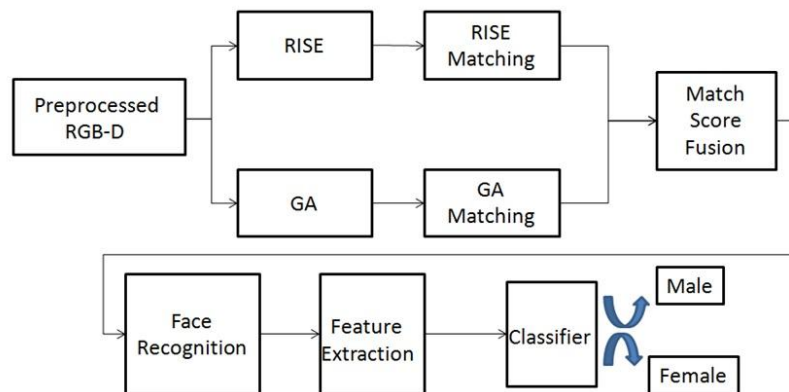


Fig.2. Illustrating the steps involved in the Face and Gender recognition.

The major modules are listed below

A. Preprocessing

First, face region need to be extracted from the RGB image for that an automatic face detector (Viola-Jones face detector) is used. From the depth map also the corresponding face region is extracted. In case if there are two faces detected in an image, then the face with maximum area will be selected.

B. Face Recognition

After preprocessing three major steps are involved in face recognition and it is explained in detail below

1) *RISE (RGB-D Image Descriptor Based on Saliency and Entropy)*: In the RISE algorithm [1], four entropy maps are computed corresponding to both RGB image and depth map. A saliency map is computed for the RGB image. The HOG features are extracted from these five entropy/saliency maps. Five HOG descriptors are concatenated and it is given as input to the trained Random Decision Forest (RDF) classifier to obtain the match score. For more information about RISE algorithm refer [1]. The steps involved in the RISE algorithm are shown in the Fig.3 and explained below.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

a) *Entropy and Saliency*: Entropy is defined as the measure of uncertainty in a random variable [12]. Entropy of an image gives the variance in the gray scale in a local neighborhood. The entropy H of an image is given in eq. (1), where x is the neighborhood taken.

$$H(x) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i) \quad \text{eq. (1)}$$

Where $p(x_i)$ gives the probability that grayscale x_i appears in the neighborhood and n denotes the total number of possible grayscale values, i.e., 255. If x is a $M_H \times N_H$ neighborhood then $p(x_i)$ can be written as follows

$$p(x_i) = \frac{n_{x_i}}{M_H \times N_H} \quad \text{eq. (2)}$$

The number of pixel in the neighborhood with value x_i is denoted by n_{x_i} and the total number of pixels in the neighborhood is taken as $M_H \times N_H$. For entropy map computation, neighborhood size is fixed at 5×5 and input RGB image is converted to grayscale before the computation.

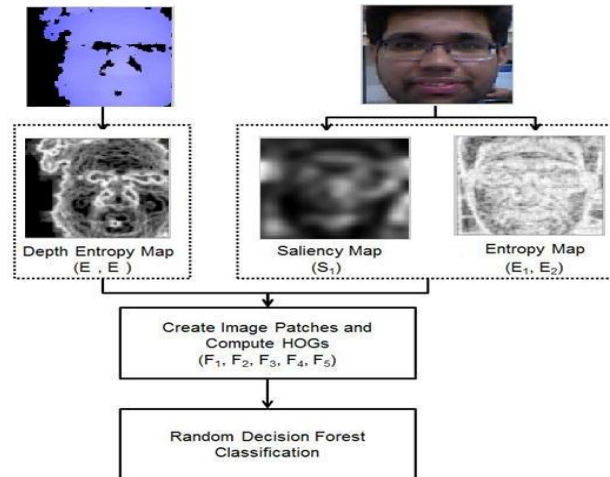


Fig.3. Illustrating the steps involved in the RISE algorithm.

Saliency of the RGB image is computed to obtain useful facial information. This work utilizes the approach proposed by Itti et al. [13]. Simple Implementation of saliency [19] by Jonathan Harel in MATLAB is used in this work. Feature filtering is done to extract colors, intensity and orientations features. Compute center-surround differences by taking the difference between a fine (center) and a coarse scale (surround) for a given feature. Center-surround difference is computed to obtain feature maps. A total of 42 feature maps are extracted from the image which includes feature maps of 6 for intensity, 12 for color and 24 for orientations. Then these feature maps are combined to obtain saliency map. The Fig.3 shows the saliency map obtained from RGB image.

After obtaining Entropy maps and Saliency map, let the RGB-D image be denoted as $[I_{rgb}, I_d]$, where I_{rgb} is the entropy map obtained from RGB image and I_d is the entropy map obtained from depth map, both of size $M \times N$. Two image patches are selected from both I_{rgb} and I_d .

Two patches P_1 and P_2 are extracted from the I_{rgb} . P_1 of size $\frac{M}{2} \times \frac{N}{2}$ and P_2 of size $\frac{3M}{4} \times \frac{3N}{4}$ centered at $[\frac{M}{2} \times \frac{N}{2}]$. Similarly two patches P_3 and P_4 are extracted from I_d . Four entropy maps $E_1 - E_4$ represent the four patches $P_1 - P_4$ taken from I_{rgb} and I_d using eq.(3):

$$E_i = P_i, \text{ where } i \in [1, 4] \quad \text{eq. (3)}$$

E_1, E_2 represents the entropy of the color image and E_3, E_4 represent the entropy of depth map.

b) *HOG*: HOG [14] descriptor produces the histogram of a given image in which pixels are binned according to the magnitude and direction of their gradients. Let HOG histogram be denoted by (\cdot) ; the RISE algorithm computes HOG of entropy maps using the equation stated below:

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

$$F_i = D(E_i), \text{ where } i \in [1, 4] \quad \text{eq. (4)}$$

Here, F_1 and F_2 represent the HOG of entropy map E_1 and E_2 of RGB image respectively. Similarly, F_3 and F_4 represent the HOG of entropy maps E_3 and E_4 of depth map respectively. F_1 and F_2 capture texture information whereas F_3 and F_4 capture the depth information. F_5 represent the HOG descriptor of saliency map S and it is given in eq.(5). By using an ordered concatenation of the five HOG histograms the final descriptor F is created as shown in eq.(6).

$$F_5 = D(S(I_{rgb})) \quad \text{eq. (5)}$$

$$F = [F_1, F_2, F_3, F_4, F_5] \quad \text{eq. (6)}$$

c) *Classification*: Random Decision Forests (RDFs) [15] classifier is used for identifying the identity of a given probe. First RDF is trained with feature vector F of each known user. While testing the identity of the probe, it will compare the feature vector F of the probe with that of stored users. Here subject identification number is taken as the class label and a class label is given to each known subject. So the number of classes will be equal to the number of subjects. Feature vector of the probe is given as input to the trained RDF which provides a match score for each class. The probability with which the feature vector belongs to a particular class is denoted by this match score.

2) *GA (Geometric Attribute)*: The Geometric attribute (GA) in the face is computed from the RGB image. To describe a face multiple geometric attributes can be utilized such as the distances between various key facial features including eyes, nose and chin. An overview of the GA approach is illustrated in Fig.4. The GA approach consists of the following steps.

a) *Keypoint Labeling*: Geometric attributes needed to be extracted from facial image. For that first a few facial key points are located in the RGB image. The points such as nose and eyes can be extracted by using Viola-jones algorithm. In a detected face image Viola-Jones algorithm uses haar features to detect nose and eyes. By making use of landmark such as eyes, the location of nose bridge is located. Chin is estimated by making use of the slope made by nose bridge and nose tip. Geometric attribute computation is done based on these landmark points.

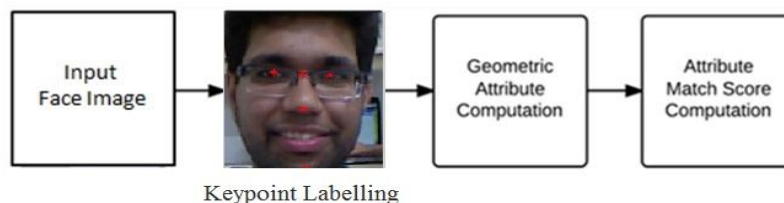


Fig.4. Steps involved in the GA approach.

b) *Geometric attribute computation*: For geometric attribute computation, distance between the landmarks points such as eyes, nose bridge, nose tip and chin need to be calculated. Some of the computed distances are listed below, that includes: Inter-eye distance, Eye to nose bridge distance, Nose bridge to nose tip distance, Nose tip to eyes distance, Nose tip to chin distance, Nose bridge to chin distance, Overall length of the face.

From multiple gallery images these parameters are measured, as these measured values may vary across pose and expression. For each gallery image, attributes are computed individually and the resulting distances are averaged. Match score is computed separately for each subject in the gallery.

c) *Attribute Match Score Computation*: The match score Φ is computed to know the similarity between each subject in the gallery and the probe. It is computed using the eq.(7).

$$\Phi = \sum_{i=1}^N w_i \times (A_i - a_i)^2 \quad \text{eq. (7)}$$

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

Here, the i^{th} attributes of the probe image and the gallery image is represented by A_i and a_i respectively. Weight of the i^{th} attribute is represented by w_i and the total number of attributes is represented by N. Once the match score calculation is completed it can be used for identification.

3) *Match Score Fusion*: The match scores obtained by RISE and GA algorithms is combined using match score level fusion. For match score level fusion, it makes use of weighted sum rule [16]. The match scores obtained by RISE approach and GA approach is represented by Φ_{RISE} and Φ_{GA} respectively. The final match score is computed as follows,

$$\Phi_{final} = w_{RISE} \times \Phi_{RISE} + w_{GA} \times \Phi_{GA} \quad \text{eq. (8)}$$

The weights assigned to the RISE and GA approach are represented as w_{RISE} and w_{GA} respectively. If Φ_{final} is greater than the predefined threshold then the user is recognized and corresponding ID of the user will be displayed based on the matching score. If Φ_{final} is lesser than the predefined threshold then the user will be rejected and a message showing unknown user will be displayed.

C. Gender Recognition

As an extension to the existing system, proposed a method to perform gender recognition after face recognition. Recognizing human gender is important since people respond differently based on gender of the person. From the Recognized face Gender Recognition is to be done and it is shown in Fig.5. For that features needed to be extracted from the face image. So HOG features [14] are extracted from the face image. The histogram of oriented gradients descriptor gives the local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions. First HOG features are obtained from the facial images of male and female. There after SVM classifier is trained with male features and then the classifier is trained with female features. That concludes the training part. For testing a facial image, First the HOG feature is computed for the image then the feature is given to the SVM classifier. Now the classifier will predict the gender of the person based on the HOG features.

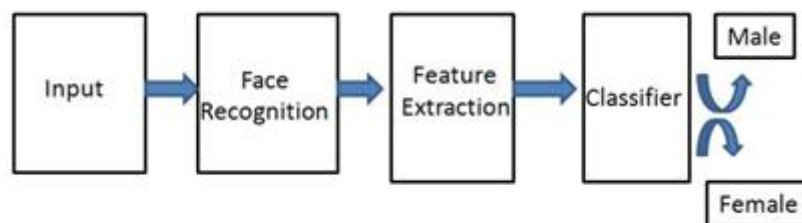


Fig.5. Illustrating the steps involved in Gender Recognition.

1) *HOG and SVM*: From the Recognized face Gender Recognition is to be done. For that features needed to be extracted from the face image. So HOG [14] features are extracted from the face image. The histogram of oriented gradients descriptor gives the local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions. The image is divided into small connected regions called cells, and for the pixels within each cell, a histogram of gradient directions is compiled. The descriptor is then the concatenation of these histograms.

Train the SVM classifier: First HOG features are obtained from the facial images of male and female. Thereafter SVM classifier [17] is trained with male features and then the classifier is trained with female features. SVM classifier is trained with male and female HOG features separately.

Testing Gender: For testing the Gender from the facial image, the HOG feature is computed for that image. Then the feature is given to the SVM classifier. Now the classifier will predict the gender of the person based on the HOG features.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

2) *DWT and SVM*: For performance evaluation gender recognition is performed using previous work which uses discrete wavelet transform (DWT) [18]. DWT extract features by decomposing image in frequency domain into subbands at different scales. Here single-level two-dimensional wavelet decomposition is done to extract features.

Train the SVM classifier: First features are obtained from the facial images of male and female, then dimensionality reduction is done using PCA (Principal component analysis). Thereafter SVM classifier is trained with male features and then the classifier is trained with female features.

Testing Gender: For testing the Gender from the facial image, the feature is computed for that image. Then the dimensionality of the feature is reduced using PCA. Thereafter it is given to the SVM classifier. Now the classifier will predict the gender of the person based on the DWT features.

IV. EXPERIMENTAL RESULTS

The experiments are performed on the IIIT-D RGB-D dataset [20] to analyze the performance of the face recognition that uses RISE and GA approach. Thereafter, the performance of gender recognition using HOG and SVM classifier is compared with DWT and SVM classifier. IIIT-D RGB-D dataset consists of 4605 RGB-D images. Gallery size is fixed at six images per subject. The proposed system is implemented with MATLAB. Sample images of different subjects taken from IIIT-D RGB-D database is shown in the Fig.6. For each subject, RGB-D images are stored in the database and an RGB-D image consists of an RGB image and depth map as shown in the figure.

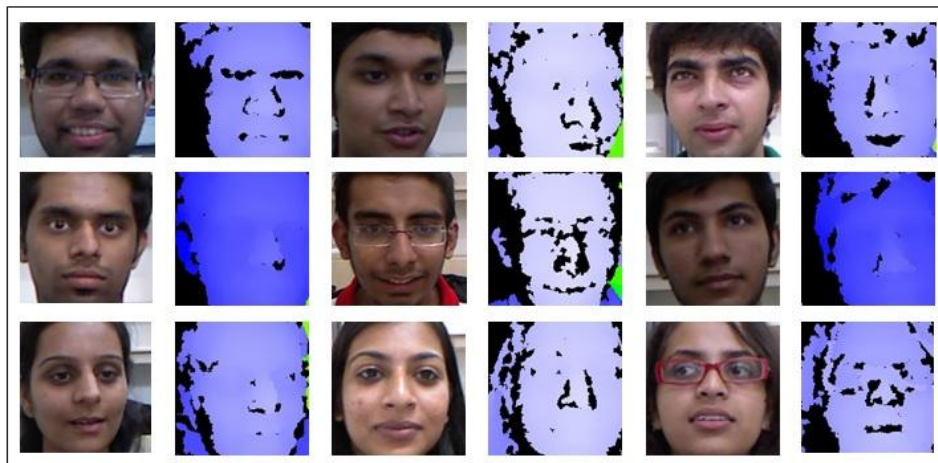


Fig.6. Sample images from the IIIT-D RGB-D database.

For testing the performance of RGB-D Face recognition, it is tested on IIIT-D RGB-D Database. RGB-D Face recognition uses RISE and GA algorithm. The number of subjects selected for training is 42. Here tested on 310 facial images of 64 subjects and got an accuracy of 89%. For testing the performance of Gender Recognition, it is tested on IIIT-D RGB-D Database. Gender recognition is done using two methods. First Gender recognition is done using HOG and SVM classifier and to compare its performance, gender recognition is also done using DWT and SVM classifier. The result obtained after performing gender recognition using HOG and DWT is shown in the Fig.7. The figure shows an example result where male and female subjects are taken for gender recognition. HOG and SVM classifier correctly recognized both male and female subjects. Similarly DWT and SVM classifier also correctly recognized both male and female subjects and it is shown in the figure.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

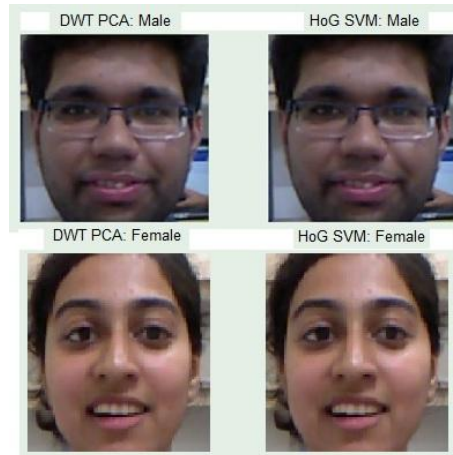


Fig.7. Gender Recognition using HOG and DWT.

First the performance of gender recognition using HOG and SVM is done. Here trained the classifier with 285 male and 253 female images. Tested the classifier with 380 images of which 226 are Male and 164 are Female. These images are taken from IIIT-D RGB-D Database. An accuracy of 91% is obtained for gender recognition using HOG and SVM. Performance of HOG and SVM on the IIIT-D RGB-D face database is plotted in the Fig.8. The performance is plotted using an ROC curve. ROC(receiver operating characteristic) curve, is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. The performance is plotted using inbuilt function for plotting ROC curve in Matlab and the input to the function is Actual output and Predicted output when performing gender recognition using HOG and SVM.

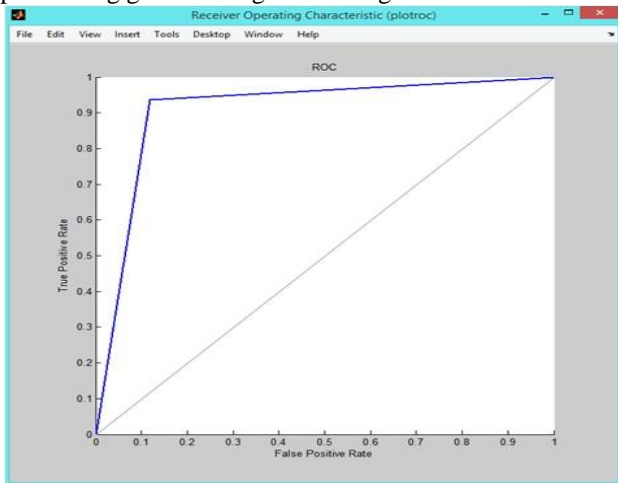


Fig.8. Performance of HOG and SVM

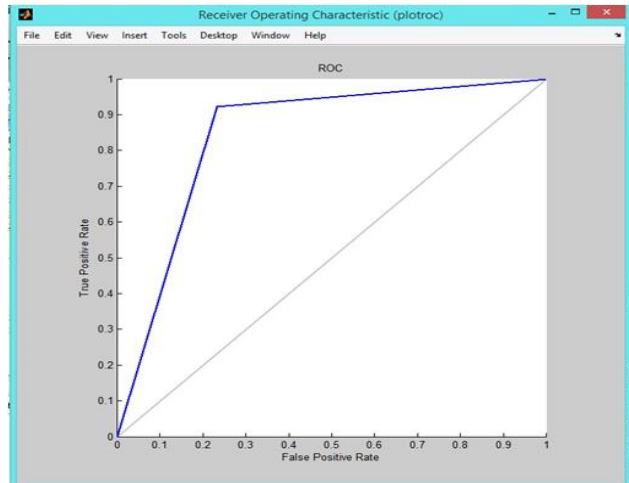


Fig.9. Performance of DWT and SVM

Now the performance of gender recognition using DWT and SVM is done. Here trained the classifier with 285 male and 253 female images. Then tested with 380 images of which 226 are Male and 164 are Female. These images are taken from IIIT-D RGB-D Database. An accuracy of 85% is obtained for gender recognition using DWT and SVM. Performance of DWT and SVM on the IIIT-D RGB-D face database is plotted in the Fig.9. The performance is plotted using an ROC curve as explained above. The curve shows that gender recognition using HOG achieves higher accuracy when compared with that of using DWT.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

Table.1. Performance Evaluation of Face and Gender Recognition

Recognition	Total Tested Images	Correctly Recognized	Incorrectly Recognized	Accuracy
Face Recognition	310	276	34	89%
Gender Using HOG	380	347	33	91%
Gender Using DWT	380	326	54	85%

The experimental results obtained for all the experiments are summarized in Table.1. For face recognition tested on 310 RGB-D facial images and by correctly recognizing 276 RGB-D facial images got an accuracy of 89%. For gender recognition using HOG and SVM classifier tested on 380 RGB facial images and by correctly recognizing 347 RGB facial images got an accuracy of 91%. For gender recognition using DWT and SVM classifier tested on 380 RGB facial images and by correctly recognizing 326 RGB facial images got an accuracy of 85%. Hence the accuracy obtained by HOG and SVM classifier is more when compared with DWT and SVM classifier.

V. CONCLUSION AND FUTURE WORK

In order to improve the performance of face recognition in the presence of covariates such as pose, expression and illumination, RGB-D image is used. GA algorithm is proposed to extract and match geometric attributes. GA is then combined with the RISE algorithm for identification. Here geometric attribute is extracted from color image as depth map contains lots of noise. In various biometric applications, gender recognition from facial images plays an important role. In order to increase the security of the system proposed a method to perform gender recognition after performing face recognition. Here investigated Histogram of oriented gradients (HOG) Descriptor for gender recognition. So HOG Features are extracted from the facial image. SVM classifier is then used to recognize the facial features for gender recognition. The experimental results indicate that the gender recognition using HOG achieves higher accuracy on color images when compared with gender recognition using DWT. Future works could include canonical frontal view obtained irrespective of their initial pose can be done for face recognition across different pose and also method that improve the accuracy of gender recognition can be included.

REFERENCES

1. G. Goswami, M. Vatsa, and R. Singh, "RGB-D Face Recognition with Texture and Attribute Features", IEEE Transactions On Information Forensics And Security, Vol. 9, No. 10, October 2014
2. Pamplona Segundo M, Sarkar S, Goldg D, Silva L, and Bellon O, "Continuous 3D Face Authentication using RGB-D Cameras," in Proc IEEE Conf. Computer Vision and Pattern Recognition Workshops, pp. 64-69, Jun. 2013.
3. B. Y. L. Li, A. S. Mian, W. Liu and A. Krishna, "Using Kinect for Face Recognition Under Varying Poses, Expressions, Illumination and Disguise," in Proc. IEEE Conf. Applications of Computer Vision (WACV), pp. 186-192, Jan. 2013.
4. C. Ciaccio, L. Wen, and G. Guo, "Face Recognition Robust to Head Pose Changes Based on the RGB-D Sensor," in Proc. IEEE Conf. Biometrics: Theory, Applications and Systems (BTAS), pp. 1-6, 2013
5. N. Erdogmus, and J. Dugelay, "3D Assisted Face Recognition: Dealing With Expression Variations," IEEE Transactions On Information Forensics And Security, Vol. 9, No. 5, May 2014
6. G. Goswami, S. Bharadwaj, M. Vatsa, and R. Singh, "On RGB-D Face Recognition using Kinect," in Proc. IEEE Conf. Biometrics: Theory, Applications and Systems (BTAS), pp. 1-6, 2013
7. O. Nikisins, K. Nasrollahi, M. Greitans and T. B. Moeslund, "RGB-D-T based Face Recognition," in Proc. IEEE 22nd Int. Conf. Pattern Recognition (ICPR), pp. 1716 - 1721, 2014.
8. F. Tsalakanidou, S. Malassiotis, and M. G. Strintzis, "Face Localization and Authentication Using Color and Depth Images," IEEE Transactions On Image Processing, Vol. 14, No. 2, February 2005.
9. N. Nourbakhsh and K. R. Safabakhsh, "3D constrained local model-based face recognition using Kinect under variant conditions," in Proc. IEEE 7th Int. Conf. Telecommunications (IST), pp. 361 - 366, 2014.
10. R. Min, N. Kose, and Jean-Luc Dugelay, "KinectFaceDB: A Kinect Database for Face Recognition," IEEE Transactions On Systems, Man, And Applications: Systems, Vol. 44, No. 11, November 2014.
11. Z. Zhang, W. Zhang, J. Liu, and X. Tang, "Multiview Facial Landmark Localization in RGB-D Images via Hierarchical Regression With Binary Patterns," IEEE Transactions On Circuits And Systems For Video Technology, Vol. 24, No. 9, September 2014.
12. A. Rnyi, "On measures of entropy and information," in Proc. Berkeley Symp. Math. Statist. Probab., pp. 547-561, 1961
13. L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. Anal. Mach. Intell., vol. 20, no. 11, pp. 1254-1259, Nov. 1998.
14. Dalal, N. and B. Triggs. "Histograms of Oriented Gradients for Human Detection," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 886-893, June 2005
15. T. K. Ho, "Random decision forests," in Proc. 3rd Int. Conf. Document Anal. Recognit., pp. 278-282, Aug. 1995
16. A. A. Ross, K. Nandakumar, and A. K. Jain, "Handbook of Multibiometrics," vol. 6. New York, NY, USA: Springer-Verlag, 2006.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

17. B. Moghaddam, M-H. Yang "Gender Classification with Support Vector Machines" IEEE International Conference on Automatic Face and Gesture Recognition (FG), pp. 306-311, March 2000
18. S. A. Khan, M. Nazir, N. Naveed and N. Riaz, "Efficient Gender Classification Methodology using DWT and PCA," IEEE Conference on Multi-topic, pp. 155-158, Dec 2011.
19. Jonathan Harel, A Saliency Implementation in MATLAB: <http://www.klab.caltech.edu/~harel/share/gbvs.php>
20. <https://research.iiitd.edu.in/groups/iab/rgbd.html>

BIOGRAPHY

Nita Anil received her B.tech degree in Department of Computer Science from Mahatma Gandhi University, Kottayam in 2013 and pursuing M.Tech in Department of Computer Science at Viswajyothi College of Engineering and Technology, Vazhakulam, Kerala, India. Her research interests are Image Processing, Biometrics etc.

Mr. Sivadas T Nair is an Assistant Professor in the Computer Science and Engineering Department, Viswajyothi College of Engineering and Technology, MG University, Kerala, India. He received M.Tech in Department of Computer Science from Mahatma Gandhi University, Kottayam in 2011. He completed his M.phil from the Madurai Kamaraj University during the year 2007-08.