A Novel Face Features Extraction and Neural Network Based Face Quality Assessment Technique

Vikram Mutneja[1*] and Satvir Singh[2]

[1]*Ph.D. Research Scholar, I.K. Gujral Punjab Technical University,
Kapurthala, Punjab (India)*
[2]*Associate Professor, I.K. Gujral Punjab Technical University,
Kapurthala, Punjab (India)*
[1]*vikram.mutneja@gmail.com, [2]drsatvir.in@gmail.com*

Abstract

Face quality assessment has evolved as eminent research area because of its significance in the task of face detection and recognition in unconstrained images from surveillance videos. It is a critical step particularly in low resolution surveillance cameras based systems in which it is important to discard the poor quality faces and select the faces having quality above a minimum acceptable threshold level so as to improve the performance of overall system. In this work, we have developed a novel neural network based system for face quality assessment. The various parameters which have been used in this work for the assessment of facial image quality are: In-Plane rotation, Out-plane rotation, Resolution, Color ratio, Sharpness, Brightness and Entropy. Face databases have been selected and their ground truth located of the facial landmarks which are: Left Eye, Right Eye, Nose Tip and Lip Center. The modified version of viola-Jones algorithm has been used to locate the faces and facial landmarks. The computed facial features have been used as inputs and the error function computed based upon the difference in estimated location of landmarks and corresponding ground truth values has been used as the output to train the neural network. The proposed technique has been found to give reliable quality assessment of the faces.

Keywords: Facial Features Extraction, Face Quality Assessment, Neural Network, Low Resolution Surveillance Videos, Face Quality Estimation

1. Introduction

As surveillance cameras and mobile cameras are nearly omnipresent now a days, processing of images and videos to extract the faces and facial features for the purpose of human identification is important requirement in most of the surveillance applications. It has been well established fact that the face quality assessment is important step for any face processing based systems, particularly in the face biometrics. The type of the technique employed for face quality assessment may be dependent upon the application areas such as face biometrics, face quality enhancement, video summarization, video compression etc. The technique employed may also be dependent upon the hardware platform such as CPU, GPU, FPGA or mobile platform. Different methods have been proposed for assessment of various factors such as blurring, lighting, head pose, facial expressions [2] for the face quality assessment.

The objective of face quality assessment is to estimate the quality of human faces in a given image or frame of a video. A lot of work has been done in the area of face quality assessment and indexing in images as well as in videos, surveyed such as by [3], [1], [4] and [16]. Most of the face quality assessment algorithms recently researched upon have

---

* Corresponding Author

been found to be either based on facial features computation or encoded features based techniques. Of these, the majority of the methods are based upon features computation, which are further used for the estimation of aggregated face quality index. Very few techniques have been found to be based upon machine learning and statistical estimation.

Researchers have used various approaches for performance evaluation of their face quality assessment technique, depending upon the target objective for which the quality assessment has been done, such as visual appearance, face recognition, resolution enhancement, spoof detection etc. The general methodology in all cases involves evaluating face quality assessment technique on relevant face databases containing facial images of quality ranging from poor to excellent based on requisite quality metrics, so as to check the results for all possible types of inputs. The output FQA (Face Quality Assessment) score should be in proportion to the quality of input face image in terms of human perception or the target biometric identification algorithm.

There are many directions in which work in this domain has leapt. If we talk about low resolution images, the faces we obtain from the images or video frames can be classified in two categories. One of the categories is the faces in which the facial landmarks can be located such as eyes, nose tip, lip centre, chin centre etc. Other are the faces which have so low resolution that facial landmarks can not be located. Obviously the quality of such facial images is very low. The type of the quality assessment technique to be deployed depends upon three factors. One is the facial features being used, secondly the accuracy of the algorithms in estimating their values, and thirdly upon the quality of input images. All of these factors are inter-dependent in a given environment where system has to be deployed but very less existing works have targeted to use this interdependency in designing the algorithm for the face quality assessment. We have encoded this interdependence in terms of using the error in localizing the facial landmarks which are left eye, right eye, nose tip and lip center, as output of the neural network for obtaining the score of face quality assessment. The modified viola-jones algorithm ([12], [17]) has been used in localizing the faces and facial landmarks eyes, nose and lips. The inputs to the neural network are the features which are In-plane angle (Roll), Out-plane angle (Yaw), Sharpness, Brightness, Resolution, Color ratio and Entropy of faces in input images.

This paper is structured as follows: Section 2 provides Literature Survey, Section 3 and 4 provides details of facial features being used for quality assessment and proposed FQA algorithm respectively. Section 5 explains experimental setup, Section 6 discusses experimental results, and finally Section 7 discusses conclusions and future Scope.

2. Literature Survey

A lot of the work in recent years has been done in the area of face quality assessment or alternatively addressed as face quality estimation. The viola jones algorithm [17] has been used to detect faces as well as localize facial landmarks. [5] worked on analysis of viola jones algorithm for the selection of its parameters to target object detection in specific conditions such as illumination, maximum search distance *etc*.

[13] questioned and investigated the existence of face quality measures. Authors interestingly used the concept of oracle for image quality in the context of face recognition by introducing greedy pruned ordering (GPO) analysis to approximate facial image quality oracle. The upper bound on quality measures on a face recognition algorithm and data set was ascertained through GPO analysis. Authors reported to having investigated many quality measures through GPO analysis such as Edge Density, Exposure Time, Eye Distance, Face Saturation, Face Standard Deviation, F stop, Focus, ISO Speed rating, Left/right hist, Illumination Direction, Mean Ratio, Pose and finally an SVM trained quality measure.

[2] studied the effect of factors blurriness, lighting conditions, head pose variations and facial expressions on quality of facial images. Authors used the eigenface technique to develop the face recognition prediction function for investigating the effect of above factors on quality of facial images and developed the classifier for images affected by facial expressions.

[9] proposed technique of features selection based on modified particle swarm optimization (MPSO) for face recognition system to search the solution space for an optimum value where features are carefully selected according to a well defined discrimination criterion. Image of the face is divided first into sub-regions, afterwards MPSO algorithm is applied to coefficients extracted by Discrete Wavelet Transform (DWT). Authors illustrated the experimental results with the minimal set of selected features using different experimental protocols on several databases, including Yale Face, FEI and ORL. Authors reported their proposed method to be robust to varying illumination, facial expressions and poses at certain angles.

[6] presented an approach for standardization of facial image quality to exclude the poor quality biometric samples from input, so as to enhance the effective performance of the biometric system. Authors worked on developing facial symmetry based methods for face quality assessment and to study the effect of non-frontal lighting and facial pose in overall face quality estimation.

[10] proposed an overall quality assessment system based on statistical learning, which is trained on the subjective quality scores, and is with a high fidelity to the human vision system (HVS) model. Authors employed a hierarchical binary decision tree classifier based on support vector machines (SVM) to categorize the overall quality of imput facial images into five levels: excellent, good, average, fair and poor. Afterwards they employed classifier fusion process to improve the performance. Authors reported the performance of their proposed system to be significantly consistent with human perception on a constructed large scale database comprising of 22720 various facial images.

3. Facial Features

Following facial features have been used as inputs to train the neural network for face quality assessment.

3.1. Head-Pose Estimation

The head-pose estimation has been used to compute the head pose orientations in terms of angles roll and yaw. The methods used for determining head pose may be categorized as local or global ([11], [7]). Global methods use template matching technique while local methods use a set of facial features like eyes, eyebrows and lips to estimate the head-pose. In this work, we have developed local methods to estimate in-plane head-pose *i.e.*, roll and out-plane head-pose *i.e.*, yaw, as formulated in equations (1) and (2).

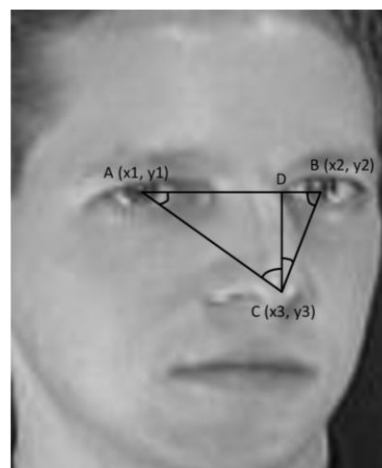1. In-Plane rotation (Roll) (refer to Figure1)

$$\Theta_r = \arctan(BC/AC) \qquad (1)$$

2. LR ratio which is measure of Out Plane rotation (Yaw) (refer to Figure 2)

$$LR\_RATIO = BC/AC \qquad (2)$$
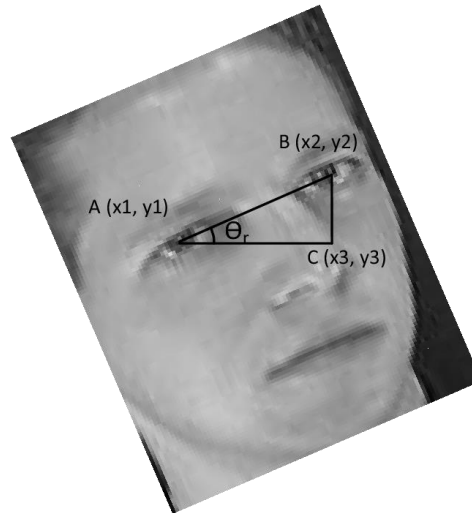
Figure 1. Head Pose: Roll

Figure 2. Head Pose: Yaw

3.2. Sharpness

Normally the images of people in real world applications captured by the surveillance cameras are affected by motion blur. The images which are well focussed have better sharpness. The performance of the face recognition algorithm is significantly affected by the sharpness of the biometric sample. Therefore the sharpness is an important factor which contributes towards the overall quality assessment of the captured facial images. We have done the estimation of the sharpness of each image to assess their blurriness by using the method as expressed in equation 3.

$$Sh(X_i) = mean|X_i - lowpass(X_i)| \tag{3}$$

where X_i is the i_{th} face image in the given dataset or video, $Sh(X_i)$ is its sharpness, low pass is a simple 3*3 mean filter.

3.3. Brightness

The accuracy of extraction of facial features is directly related to the brightness of the input facial image. Therefore we have used brightness as significant component in face

quality assessnment. The overall brightness of input facial image has been computed as a mean of illumination component of the face in YCbCr color space.

3.4. Resolution

The surveillance cameras are placed far away from the location in order to capture wide area. As a result, the size of the faces in frames from surveillance videos is usually very small. We have used our previous work of modified viola-jones face detection algorithm [12] to detect the faces. The resolution has been measured as the the size of the detected faces by our face detection algorithm.

3.5. Color Ratio

This is measure of number of colors present in the facial image with respect to the overall color range in case of color image, and mesaure of number of intensity levels present in case of gray scale image, mathematically formulated as per equations 4 and 5 respectively.

For Color Image

$$C_R = U_c / (256)^3 \tag{4}$$

For Gray scale Image

$$C_R = U_c / (256) \tag{5}$$

Where C_R is color ratio and U_c is the number of unique colors or intensity levels present in input facial image.

3.6. Entropy

It is computed as statistical measure of degree of randomness in input facial image. In our work, we have computed it as per equation (6).

$$E = -sum(N.*log_2(N)) \tag{6}$$

where N is the number of histogram counts in the histogram of input facial image.

4. Proposed FQA Algorithm

In the proposed FQA algorithm, we have developed a neural network based system which is trained to give reliable face quality assessment score. The process of training of features in the proposed neural network algorithm requires selection of a dataset of faces of varying quality in terms of their resolution, pose, brightness and their overall quality. Four face databases have been used in training of our neural network, which are ScFace [8], FERET, ATT_faces Faces and Caltech. The ScFace database has been selected because it contains low resolution facial images from surveillance videos with pose. Portions of the research in this paper use the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office [15], [14]. FERET dataset contains facial images of varying size. However their resolution and quality is better as compared to that of SCFace dataset. ATT_faces database containing small size faces with slight poses has been downloaded from webpage of AT&T lab, Cambridge university and fourth database of faces was collected by Markus Weber at California Institute of Technology, containing frontal faces of good quality of 27 people. The ground truth of facial landmarks are created manually in last two datasets while, it were available in former two databases.

Following steps have been carried in the proposed FQA algorithm.

1. Data sets were downloaded containing the training images and ground truth of following facial landmarks.

   - Eyes

   - Nose Tip

   - Lips

   These datasets have been selected based upon the range of quality of images required for training of neural network based system. Training images are selected randomly from these datasets and a new dataset containig 400 images (100 images from each database) is created. This combined dataset is being used as training dataset to train the neural network based system.

2. Facial features are extracted from the faces which are used as inputs to the neural network based system. These features have been listed in previous section.

3. Error function is computed which is used as output of the NN based system. Error function is computed of facial features extraction system with respect to the ground truth values of all the four major facial landmarks which are Left eye, Right eye, Nose tip and Lips center. If (x, y) are co-ordinates of a particular facial landmark as calculated by the algorithm, and (xg, yg) are co-ordinates of respective ground truth, respective distance D is is computed as per equation (7).

$$D_i = sqrt\,[(x-xg)^2 + (y-yg)^2\,] \tag{7}$$

Where i=1,2,3,4 corresponding to four facial landmarks respectively.

4. Maximum value of the error is the diagonal distance D of the extracted face. As there are four facial landmarks being used in calculation, let the errors calculated be denoted as D_1, D_2, D_3 and D_4. So the aggregated error function is computed as

$$Err = (D_1 + D_2 + D_3 + D_4)/(4*D) \tag{8}$$

At maximum $Err = 1$

5. Experimental Setup

The machine running windows 8.1 (64 Bit) on Intel core i3 1.9 GHz and having NVIDIA graphics processing unit GeForce GT 740M has been used to implement and test the proposed method. The work has been done in MATLAB version 8.5.0.197613 (R2015a). The face detection module has been implemented in CUDA under MATLAB environment. The CUDA files are compiled to generate the ptx code files by CUDA compiler of release 6.0, V6.0.1 with the architecture support corresponding to compute factor 3.5. The ptx code is used to create and load the CUDA kernel object in MATLAB workspace.
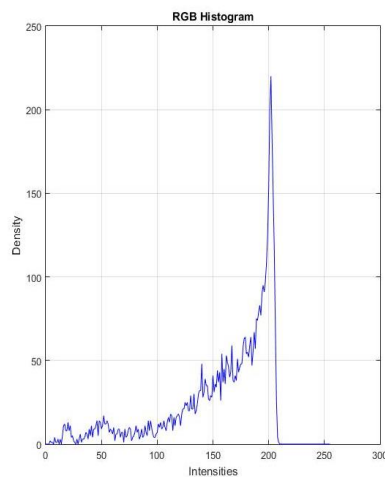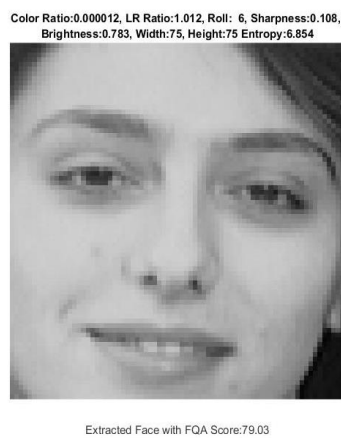
6. Results and Discussion

The proposed face quality assessment algorithm has been tested on 100 test images taken from each of the four datasets (SCface, color FERET, ATT_faces and CALTECH). The images taken are different from the training sets. We have created four folders each containing 100 images for purpose of testing. The results of applying the proposed algorithm on a random image taken from each of the dataset are shown in Figures 3, 4, 5 and 6 respectively.
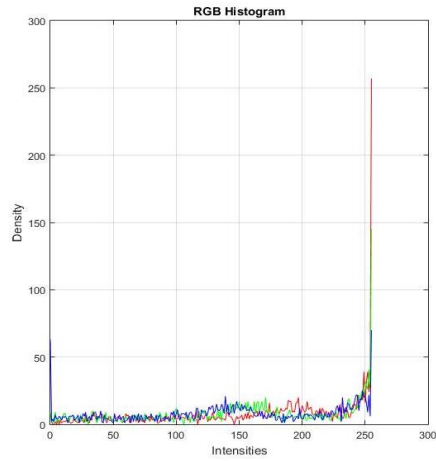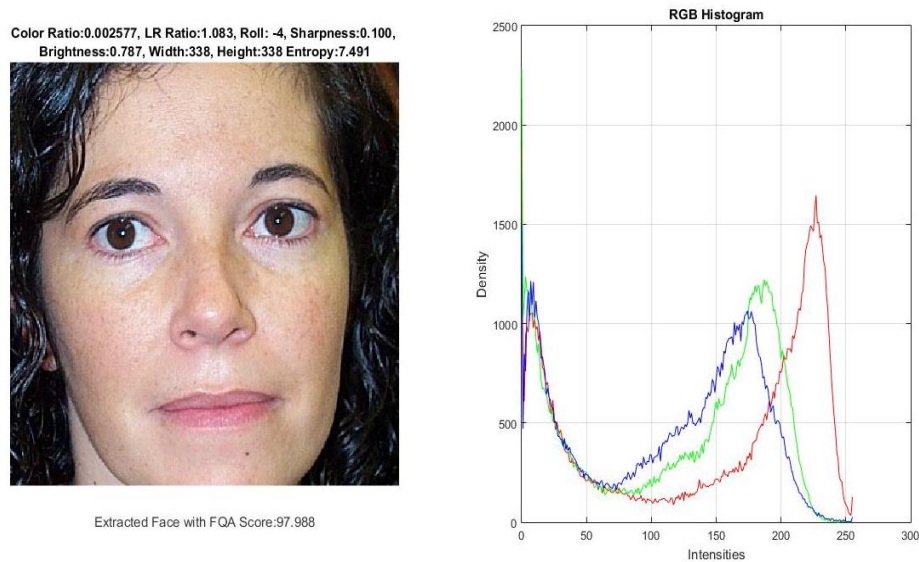
Figure 3. Result of Applying the Proposed Algorithm on a Random Image from SCface Database
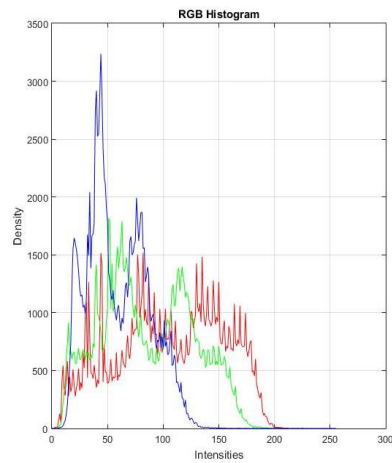
Figure 4. Result of Applying the Proposed Algorithm on a Random Image from FERET Database


Figure 5. Result of Applying the Proposed Algorithm on a Random Image from AT&T Database

Color Ratio:0.002577, LR Ratio:1.083, Roll: -4, Sharpness:0.100, Brightness:0.787, Width:338, Height:338 Entropy:7.491

Extracted Face with FQA Score:97.988

RGB Histogram

Figure 6. Result of Applying the Proposed Algorithm on a Random Image from CALTECH Database

The results of applying the proposed algorithm on all four datasets are shown in Figure 7. The results show that the FQA score computed by the proposed method are in accordance with the human visual perception which in turn is in accordance with the performance of the face recognition algorithm.
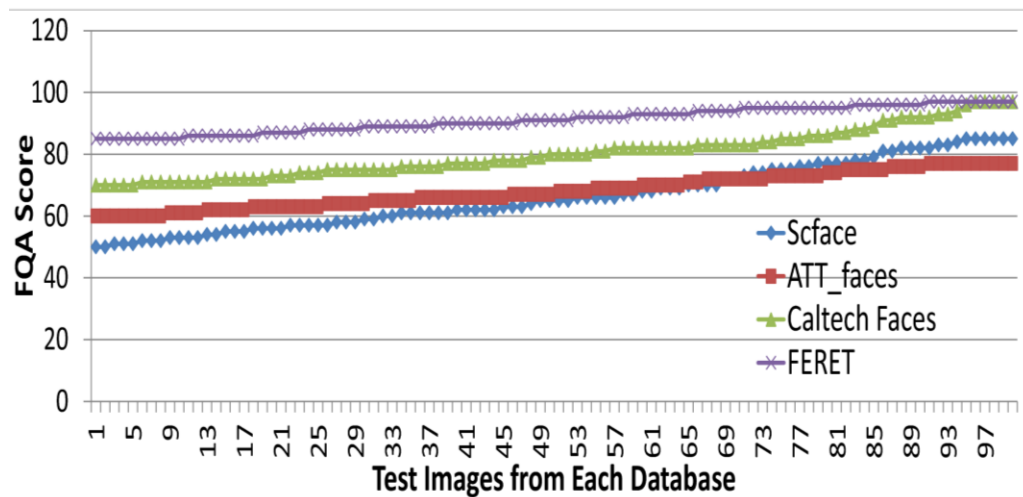
Figure 7. Comparison FQA Algorithm on All Four Datasets

7. Conclusions and Future Scope

The face quality assessment is an important step in the face biometric based systems particularly in the applications based on processing of videos from low resolution surveillance cameras based systems. The quality assessment of input facial images helps in important decision making whether to discard, perform quality enhancement or accept it as it is, so as to be processed by subsequent face identification module of surveillance cameras based systems. In this work, we have targeted the objective of developing a novel face quality assessment system. We have designed and trained the neural network based system for this task. The system is trained with inputs which are facial features: Head Pose in terms or roll and yaw, sharpness, brightness, resolution, color ratio and entropy of

the input facial images. The error function computed as difference in localization of computed facial landmarks left eye, right eye, nose tip and lip center with respect to their ground truth values is being used as output to train the neural network. Face datasets for training as well as testing are created from the databases SCface, color FERET, att_faces and Caltech, each comprising of total of 400 images, by taking 100 separate images from each of the databases. From the results of applying the proposed algorithm, it is concluded that proposed system is very effective in quality assessment of input facial images.

We propose following future directions in which the proposed work may be extended:

1. Using more of the features to measure quality of faces as inputs which may belong to category of image as well as facial part such as skin color, face shape and other photometric image quality parameters.

2. Handling of other head pose orientations in terms of pitch and severe head pose orientations.

3. Optimization of the neural network with one or multiple evolutionary techniques such as genetic algorithm, cuckoo search etc.

4. Integrating the proposed technique with video based techniques such as motion segmentation, background subtraction etc. for search space reduction and development of integrated system for processing of surveillance videos.

References

[1]   M. A. Harrison Abaza, T. Bourlai and A. Ross, "Design and evaluation of photometric image quality measures for effective face recognition", IET Biometrics, vol. 3, no. 4, **(2014)**, pp. 314-324.
[2]   M. Abdel-Mottaleb and M. H. Mahoor, "Application notes-algorithms for assessing the quality of facial images", IEEE Computational Intelligence Magazine, vol. 2, no. 2, **(2007)**, pp. 10-17.
[3]   S. Bharadwaj, M. Vatsa and R. Singh, "Biometric quality: a review of fingerprint, iris, and face", EURASIP Journal on Image and Video Processing, vol. 1, no. 34, **(2014)**.
[4]   O. Y. G. Castillo, "Survey about facial image quality", Fraunhofer Institute for Computer Graphics Research, **(2005)**, pp. 10-15.
[5]   A. Shtanko Egorov and P.Minin. "Selection of viola–jones algorithm parameters for specific conditions", Bulletin of the Lebedev Physics Institute, vol. 42, no. 8, **(2015)**, pp. 244-248.
[6]   X. Gao, S. Z. Li, R. Liu and P. Zhang, "Standardization of face image sample quality", In International Conference on Biometrics, Springer, **(2007)**, pp. 242-251.
[7]   N. Gourier, J. Maisonnasse, D. Hall and J. L. Crowley, "Head pose estimation on low resolution images", In International Evaluation Workshop on Classification of Events, Activities and Relationships, Springer, **(2006)**, pp. 270-280.
[8]   M. Grgic, K. Delac and S. Grgic, "Scface–surveillance cameras face database", Multimedia tools and applications, vol. 51, no. 3, **(2011)**, pp. 863-879.
[9]   T. Khadhraoui, S. Ktata, F. Benzarti and H. Amiri, "Features selection based on modified pso algorithm for 2d face recognition", In 2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV), IEEE, **(2016)**, pp. 99-104.
[10]   P. Liao, H. Lin, P. Zeng, S. Bai, H. Ma and S. Ding, "Facial image quality assessment based on support vector machines", In Biomedical Engineering and Biotechnology (iCBEB), 2012 International Conference on, IEEE, **(2012)**, pp. 810-813.
[11]   E. Murphy-Chutorian and M. M. Trivedi, "Head pose estimation in computer vision: A survey", IEEE transactions on pattern analysis and machine intelligence, vol. 31, no. 4, **(2009)**, pp. 607-626.
[12]   V. Mutneja and S. Singh, "Modified viola–jones algorithm with gpu accelerated training and parallelized skin color filtering-based face detection", Journal of Real-Time Image Processing, **(2017)**, pp. 1-21.
[13]   P. J. Phillips, J. R. Beveridge, D. S. Bolme, B. A. Draper, G. H. Givens, Y. M. Lui, S. Cheng, M. N. Teli and H. Zhang, "On the existence of face quality measures", In Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference, IEEE, **(2013)**, pp. 1-8.
[14]   P. J. Phillips, H. Moon, S. A. Rizvi and P. J. Rauss, "The feret evaluation methodology for face-recognition algorithms", IEEE Transactions on pattern analysis and machine intelligence, vol. 22, no. 10, **(2000)**, pp. 1090-1104.

[15] P. J. Phillips, H. Wechsler, J. Huang and P. J. Rauss, "The feret database and evaluation procedure for face-recognition algorithms", Image and vision computing, vol. 16, no. 5, pp. 295-306, **(1998)**.
[16] P. M. Shende, M. V. Sarode and M. M. Ghonge, "A survey based on fingerprint, face and iris biometric recognition system, image quality assessment and fake biometric", International Journal of Science, Engineering and Computer Technology, vol. 4, no. 4, pp. 129, 2014.
[17] P. Viola and M. J. Jones, "Robust real-time face detection", International journal of computer vision, vol. 57, no. 2, **(2004)**, pp. 137-154.

Authors

Vikram Mutneja, Mr. Vikram Mutneja is a Ph.D. research scholar in I.K. Gujral Punjab Technical University, Kapurthala (Punajb, India). He received his Bachelor's degree (B.Tech.) from Guru Nanak Dev University, Amritsar, Punjab (India) with specialization in Electronics & Communication Engineering in year 1998, Master's degree (M.Tech.) from Giani Zail Singh College of Engineering & Technology, Bathinda (Punjab) (India) under Punjab Technical University (PTU, Jalandhar) with first division in Electronics & Communication Engineering in year 2008. He is a Life Member of ISTE. He has industrial experience of around 5 years (1998-2003) in which he worked in diverse fields such as Hardware and Networking Technical Support, Software Development for Web and Embedded Systems. During his teaching experience of around 13 years (2003-2016) he worked mainly in the areas of embedded systems, VLSI and Digital Signal Processing. His current field of research includes areas of image processing for detection and extraction of facial features from low resolution surveillance videos, and acceleration using GPU computing.

Dr. Satvir Singh is Associate Professor in IKG Punjab Technical University, Kapurthala (Punajb, India). He received his Bachelor's degree (B.Tech.) from Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab (India) with specialization in Electronics & Communication Engineering in year 1998, Master's degree (M.E.) from Delhi Technological University (Formerly, Delhi College of Engineering), Delhi (India) with distinction in Electronics & Communication Engineering in year 2000 and Doctoral degree (Ph.D.) from Maharshi Dayanand University, Rohtak, Haryana (India) in year 2011. He is a Life Member of ISTE and Corporate Member of IETE. During his 15 years of teaching experience he served as Assistant Professor and Head, Department of Electronics & Communication Engineering at BRCM College of Engineering & Technology, Bahal, (Bhiwani) Haryana, India and as Associate Professor & Head, Department of Electronics & Communication Engineering at Shaheed Bhagat Singh State Technical Campus (Formerly, SBS College of Engineering & Technology), Ferozepur Punjab, India. His fields of special interest include Evolutionary Algorithms, High Performance Computing, Type-1 & Type-2 Fuzzy Logic Systems, Wireless Sensor Networks and Artificial Neural Networks for solving engineering problems.