



Towards nonuniform illumination face enhancement via adaptive contrast stretching

Aouache Mustapha¹ · Adel Oulefki¹ ·
Messaoud Bengherabi¹ · Elhocine Boutellaa¹ ·
Mustafa Almahdi Algaet²

Received: 13 November 2016 / Revised: 21 March 2017 / Accepted: 29 March 2017 /

Published online: 7 April 2017

© Springer Science+Business Media New York 2017

Abstract A face enhancement has the potential to play an important part in providing satisfactory and vast information to the face recognition performance. Therefore, a new approach for nonuniform illumination face enhancements (NIFE) was proposed by designing an adaptive contrast-stretching (ACS) filter. In a more objective manner of achieving this, an investigation usage of CS function with adjustable factors value to summarise its influence on the NIFE is examined firstly. Secondly, describe a new strategy to cater for CS adaptive factors prediction using training and testing phases. A dispersion versus location (DL) descriptor was examined in the training phase to generate the faces feature vectors. Subsequently, a frame differencing module (FDM) was developed for faces label generations. In the testing phase, the approach was examined to recognise the DL descriptor and predict face label based vocabulary tree model (VTM). Thirdly, the VTM performance was examined by referring to the area under curve (AUC) score from the receiver operating characteristic (ROC). The face quality measurement was evaluated via blind reference based

✉ Aouache Mustapha
maouache@cdta.dz

Adel Oulefki
aoulefki@cdta.dz

Messaoud Bengherabi
mbengherabi@cdta.dz

Elhocine Boutellaa
eboutellaa@cdta.dz

Mustafa Almahdi Algaet
malgaet@gmail.com

¹ Division Télécom, Centre de Développement des Technologies Avancées - CDTA, 16303, Baba Hassen, Algiers, Algeria

² Department of Computer System and Communication, Faculty of Information and Communication, Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Melaka, Malaysia

statistical measures (BR-SM), blind reference based DL-descriptors (BR-DL) and visual interpretation of the resulting images. The BR-SM performed through calculating the EME (Measure of Enhancement), EEME (Measure of Enhancement by Entropy), SDME (Second Derivative like Measure of Enhancement), SHP (Coefficient of Sharpness) and CPP (Contrast per Pixel). In addition, by using DL scatter, the BR-DL handles the specific relationship with regards to the local contrast to local brightness within the resulting face images. Four face image databases, namely Extended Yale B, Mobio, Feret and CMU-PIE were used. The final results attained prove that compared to the state-of-the-art methods, the proposed ACS filter implementation is the most excellent choice in terms of contrast and nonuniform illumination adjustment as well as providing images of satisfactory quality. In short, the benefits attained proves that ACS is driven with a profitable enhancement rate in providing tremendous detail concerning face recognition systems.

Keywords Face image · Contrast-stretching · DL descriptor · VTM · ROC · Quality measurement

1 Introduction

1.1 Motivation and research objectives

Image enhancement is primarily aimed to raise the contrast of an image with a new dynamic range and bring out the hidden image details. The recognition of human faces under varying lighting conditions is a considerably challenging issue. Situations when face images are captured in exceedingly high dynamic range scenes issue becomes increasingly difficult. A majority of the automatic face recognition systems have an assumption that images are captured under well-controlled illumination. Such a constrained condition simplifies the face segmentation and recognition. However, a surveillance system that is installed at any location will make it impossible for illumination control. It is not possible to obtain a satisfactory rate of recognition in the absence of compensation for uneven illumination. In an effort to solve the aforementioned issues, brightness-preserving uniform illumination based techniques have been examined. This is done by creating an integrated system that compensates uneven illumination through local contrast enhancement via adaptive contrast stretching (ACS) approach.

1.2 Related work

A few strategies have been suggested to make face recognition more reliable [33]. However, a majority of this approaches were primarily built on databases with face images collected under controlled illumination conditions, thus they have difficulty in dealing with variations of illumination. Here, this study reviews several current approaches for illumination invariant for image and face enhancement.

1.2.1 Enhancement based conventional approaches

One of the most popular schemes utilised for image enhancement is histogram equalization (HE), due to its speed as well as simplicity of implementation [33]. An enhanced image by HE has a significantly changed brightness; therefore the output image is saturated with either extremely dark or bright intensity values and is clouded by several artefacts and

unnatural enhancement [1, 2]. Contrast Limited Adaptive Histogram Equalization (CLAHE) [30, 47] method uses the concepts of AHE and CLHE. It uses tiles, which are small contextual regions and the peaks of the histogram are clipped. It eliminates noise amplification and increases the local contrast of the image. The disadvantage of CLAHE method is it gives unsatisfactory results when there is unbalanced contrast and increased brightness [39]. Thus, here Enhanced CLAHE method is proposed by combining CLAHE with threshold technique.

Illumination invariant face recognition was developed by Tan and Triggs [38] by utilising existing robust illumination normalisation, distance transform based matching, local texture-based face representations, multiple feature fusion and kernel-based feature extraction. The method developed by them managed to accomplish state-of-the-art recognition rates on three face databases, which were commonly utilised for assessing recognition under bad illumination conditions. In an effort to enhance the performance of methods of face recognition, Heusch et al. [15] studied the face recognition problem under changing illumination conditions. Huge improvements in classification rates were achieved by their experiments. The results of their experiments were also comparable to other state-of-the-art face recognition methods.

In face of the weaknesses of the frequently utilised illumination normalisation methods, Struc and Pavesic [44] worked on two novel methods for illumination invariant face recognition. They showed that their techniques on the Extended Yale-B database were highly promising. The Retinex method for illumination invariant face recognition was suggested by Park et al. [28]. According to the famous Retinex theory, illumination is usually estimated and normalised by pre-smoothing the input image and then dividing the estimate by the original input image. Their approach is tested on AR face database, Yale face database B and CMU PIE databases. The results of the test were encouraging and consistent, even when illuminated face images were utilised as a training set.

To compensate for illumination variations in face images, a novel image pre-processing algorithm was proposed by Gross and Brajovic [13]. No training steps, reflective surface models or knowledge of 3-D face models were involved in their algorithm. Large enhancements in performance over several standard face recognition algorithms across multiple public face databases were demonstrated. For the purpose of illumination invariant face recognition, a logarithm discrete cosine transform (DCT) was proposed by Chen et al. [8]. In their method, a small number of low-frequency DCT coefficients were removed and the inverse DCT were taken to form illumination invariant faces. However, the action of setting low-frequency DCT coefficients to zero may cause certain important features to be eliminated.

Nonlinear filtering of signals that products or a convolution of components were developed by Oppenheim et al. [27]. Their method was applied to image enhancement and audio dynamic range compression and expansion and produced very promising results. Simultaneously dynamic range compression or colour consistency was achieved when Jabson et al. [17] extended a single-scale centre/surround Retinex to a multiscale version. They also worked on a novel colour restoration approach that managed to overcome existing deficiency at the cost of a slight dilution in the consistency of the colour.

Recently, a multiscale logarithm difference edge map method was developed by Lai et al. [20] for the purpose of face recognition under changing lighting conditions. In their method, the logarithm transform was taken to convert the multiplication of surface albedo and light intensity into an addition. Subsequently, light intensity was eliminated by subtracting two neighbour pixels. Multiple multiscale edge maps were generated. They were multiplied by a weight and then all weighted edge maps were integrated to generate a robust face

map. A novel illumination invariant face recognition method via dual-tree complex wavelet transform (DT-CWT) was proposed by Hu [16]. In this method, the DT-CWT is first used to extract edges from face images. The DT-CWT coefficients are then denoised to get the multi-scale illumination invariant structures in the logarithm domain. The final step involves combining the logarithm features and extracted edges.

By utilising multi-resolution local binary pattern fusion and local phase quantization, Nikan and Ahmadi [25] studied the local gradient-based illumination invariant face recognition approach. Performance enhancement of their new method was demonstrated from their experimental results under poor illumination condition. Face recognition under changing illumination was also studied by Faraji and Qi [11]. In their method, adaptive homomorphic filtering was first done to decrease illumination effects. Next, the filtered image was stretched using an interpolative enhancement scheme. Finally, it formed eight directional edge images and generates an illumination-insensitive face map from the eight edge images.

Poddar et al. [31] suggested a Non-parametric Modified Histogram Equalization for Contrast Enhancement (NMHE), which aimed firstly to remove spikes from the histogram of the input image, clips and normalizes the result, computes the summed deviation of this intermediate modified histogram from the uniform histogram and uses this as a weighting factor to construct a final modified histogram that is a weighted mean of the modified histogram and the uniform histogram. Then, by utilising the DF of this modified histogram as the transformation function, contrast enhancement is achieved.

Similar to the mentioned techniques, Arriaga Garcia et al. in [5], introduced Bi-Histogram Equalization with adaptive sigmoid functions (BEASF) method to achieve well-enhancement by using the brightness preservation in combination with robustness and noise tolerance. This was done by first splitting the image histogram into two sub-histograms (applying the mean as a threshold) and using with two smooth sigmoids (with their origins placed on the median of each sub-histogram) to replace their cumulative distribution functions. Later, an HE and a histogram stretching within their own limits is performed, to produce a smooth and continuous mapping curve.

1.2.2 Enhancement based expert system

An expert system such as fuzzy image enhancement is founded on the idea of grey level mapping into membership function in order to form an image that has higher contrast than the original image [14]. The promising result of the fuzzy logic image enhancement accommodates some drawbacks of classical image enhancement techniques. It has the capacity to handle uncertain and vague information. In the fuzzy representation of an image, each pixel is represented by fuzzy rules and a membership function [42]. There are several image enhancement techniques base on fuzzy set theory.

Minimization of fuzziness or contrast intensification operator is probably the first fuzzy method that reduces the amount of image fuzziness to enhance the image [41]. The distance of all gray-levels to FEV is calculated in the equalization using fuzzy expected value, in order to enhance the quality of the image. Histogram equalization maximises the image information by utilising the entropy as a measure of information. Fuzzy histogram hyberbolization is the modification of histogram equalization method using suitable membership function in a logarithmic way instead of flattening it [41]. In a rule-based contrast enhancement, the parameters of the inference system are initialized, grey levels are fuzzified, inference procedure is evaluated and finally the enhanced image is achieved by the defuzzification [41]. Fuzzy relation is another image enhancement technique assuming that there is a fuzzy relation between the original and the enhanced image [41].

Enhancement of images under different illumination also adopted through expert knowledge. Therefore, Brightness Preserving Dynamic Fuzzy Histogram Equalization (BPDFHE), is another method suggested by Sheet et al. in [34]. The BPDFHE operates the image histogram in order to no remapping of the histogram peaks appears; the only redistribution of the gray-level values in the valley portions between two consecutive peaks is performed. Image enhancement using BPDFHE gain well-enhanced contrast and a few artefacts take place. But this method suffers from the drawback of requiring complicated algorithms and high computation time.

An enhanced strategy for 3D head tracking under varying illumination conditions is proposed by La Cascia et al. in [19]. The head is modelled as a texture mapped cylinder. Tracking is formulated as an image registration problem in the cylinder's texture map image. The resulting dynamic texture map provides a stabilised view of the face that can be used as input to many existing 2D techniques for face recognition, facial expressions analysis, lip reading, and eye tracking. Additionally, Choi et al. in [9] proposed an illumination-reduced feature learning method using the deep convolutional neural network (DCNN). Their learning method is mainly comprised of following two-steps: 1) learning illumination patterns for eliminating illumination effect and 2) learning for maximising the discriminative power of feature representation. Experimental results on CMU Multi-PIE database have been demonstrated that the proposed method outperforms the previous works in terms of FR accuracy.

A summary of existing enhancement filters and its application in recognising human faces shows the issue of filtering has been generally studied. However, this still represents a difficult area of research and the necessity to run more investigations to get the best images with better contrast and uniform illumination.

1.3 Contributions

In this paper, an adaptive contrast stretching (ACS) filter has been proposed for nonuniform illumination face enhancement to afford a plenty and satisfactory information to the automatic face recognition systems. The behaviour of the CS function with variable factors are investigated, a DL descriptor and frame differencing module are introduced to the VTM for face labels prediction. Additionally, the quality of filtered face images is evaluated via BR-SM, BR-DL and visual interpretation approaches. In this experiment, the enhanced face images by the ACS are compared to well-known state-of-the-art techniques like CLAHE, HE, BPDFHE, NMHE, WA and BEASF techniques.

The organisation of this paper is as such: Section 1 shows the motivation and objective along with a brief review of related works. The discussion of the proposed enhancement method that includes contrast-stretching function, face image databases, DL descriptor, VTM and adaptive filter performance evaluation is contained in Section 2. Section 3 contains experimental results and discussion. To summarise the overall enhancement performance, concise conclusions are given in Section 4.

2 Methods

2.1 Database description

In this research, face images from four publicly available databases are selected in our experiment as follows:

First, the extended Yale Face Database B (EYale-B) in [21] has 16128 images of 38 human subjects captured under 64 illumination conditions and 9 poses. The database extends the original one [12] by involving 28 new subjects. Five different light sources positioned at (12° , 25° , 50° , 77° and 90°) of the camera axis are employed for capturing the faces under illumination variation. An image with ambient (background) illumination was also captured for each subject in a particular pose. A sample of face images for one subject in the Extended Yale B database is shown in Fig. 1.

Second, the Mobio database [23] as shown in Fig. 2 contains talking face videos of 152 people. The gender statistics of the database participants are 100 males and 52 females. The database recording was performed using two types of mobile devices: laptop computers (2008 standard MacBook) and mobile phones (NOKIA N93i). The database was collected in a period of about 2 years from August 2008 until July 2010 at six distinct university sites. The recording environment was uncontrolled involving variations in illumination, facial expression and face pose. The data of each individual were captured in



Fig. 1 A sample of cropped face images for one subject in the Extended Yale-B database



Fig. 2 A sample of cropped face images for one subject in the Mobio database

12 sessions to maximise the intra-person variability. In our work, we use the facial still image the Mobio databases. The images are extracted from each video by the database collectors.

Third, the facial images of Feret database [29] were collected through 15 different sessions between August 1993 and July 1996, with every session lasting for one or two days. The database involved 1199 individuals. For each individual 5 to 11 images were collected. In total, the database contains images 14,126 face images. The images present different challenges, including person ageing, facial expressions and illumination change. A sample of face images for one subject in the Feret database is shown in Fig. 3.

Fourth, the CMU Pose, Illumination, and Expression (CMU-PIE) Database [36] have over 40,000 facial images of 68 people. For each person, many images are captured under 43 different illumination conditions, 4 different expressions and 13 different poses. The database was collected over a period of three months. The recording was performed in a dedicated room using 13 high-quality Sony DXC 9000 cameras. The images are coloured ones with a resolution of 640×486 . The illumination has been generated using a flash system of 21 different flashes. A sample of face images for one subject in the CMU-PIE database is shown in Fig. 4.

2.1.1 Contrast-stretching (CS) function

To support the localisation of key components of the face images and their characteristics, the CS function has been investigated and applied. The histogram range is stretched by the



Fig. 3 A sample of cropped face images for different subject in the Feret database

CS transformation to fill the whole intensity domain of the image. The filtered image using CS function is as follows:

$$O(i, j) = \frac{1}{1 + (G/I(i, j))^E} \quad (1)$$

where, the intensity value of the output and input images are represented by $O(i, j)$ and $I(i, j)$, respectively; the positive factors E and G control the slope function and switches the dark pixel to bright, respectively.

2.1.2 DL descriptor features

With location (L) measured by midrange and dispersion (D) measured by the range, the DL descriptor is taken as a feature extraction approach that manages the relationship between local brightness in the face images and local contrast via the utilisation of dispersion and location estimators [35]. The range (the max minus min) and midrange values (the average of min and max) utilised for the location and dispersion estimator, respectively, are the basis of these statistical measurements. Reasonable information of the image is represented by the scatter plot, which is a DL plane of dispersion versus location [24, 32]. Following the purpose of classifying the points in the scatter plots, the DL plots, have points restricted to the triangle with vertices (1, 0), (0, 0), and (0.5, 1). As shown in Fig. 5, the DL plots are subdivided into four triangles D, C, B and A with the description as indicated in Table 1. Where, HC/ML (High-Contrast / Medium-Luminance), LC/LL (Low-Contrast / Low-Luminance), MC/ML (Medium-Contrast / Medium-Luminance), and LC/HL (Low-Contrast / High-Luminance).



Fig. 4 A sample of face images for one subject in the CMU-PIE database

Using experimental and visually determined linear transformation, the [midrange, and range] pair within image is computed using (2) by employing an optimal window size of 15×15 [24, 32].

$$\begin{bmatrix} Midrange \\ Range \end{bmatrix}^T = [\min \max] \times \begin{bmatrix} 1/2 & 1/2 \\ -1 & 1 \end{bmatrix} \tag{2}$$

Figure 5a, b illustrates the behaviour examination of transfer function mapping (TFM) and DL descriptor distribution of CS function with variable values pair (E, G) on the input/output image, where the DL distribution and the transfer function map (TFM) represents the scatter plot of $O(i, j)$ versus $I(i, j)$.

Table 1 DL feature description

Description	Triangles vertex
HC/ML	$A \in \{(0.25, 0.5), (0.5, 1), (0.75, 0.5)\}$
LC/LL	$B \in \{(0, 0), (0.25, 0.5), (0.5, 0)\}$
MC/ML	$C \in \{(0.25, 0.5), (0.5, 0), (0.75, 0.5)\}$
LC/HL	$D \in \{(0.5, 0), (0.75, 0.5), (1, 0)\}$

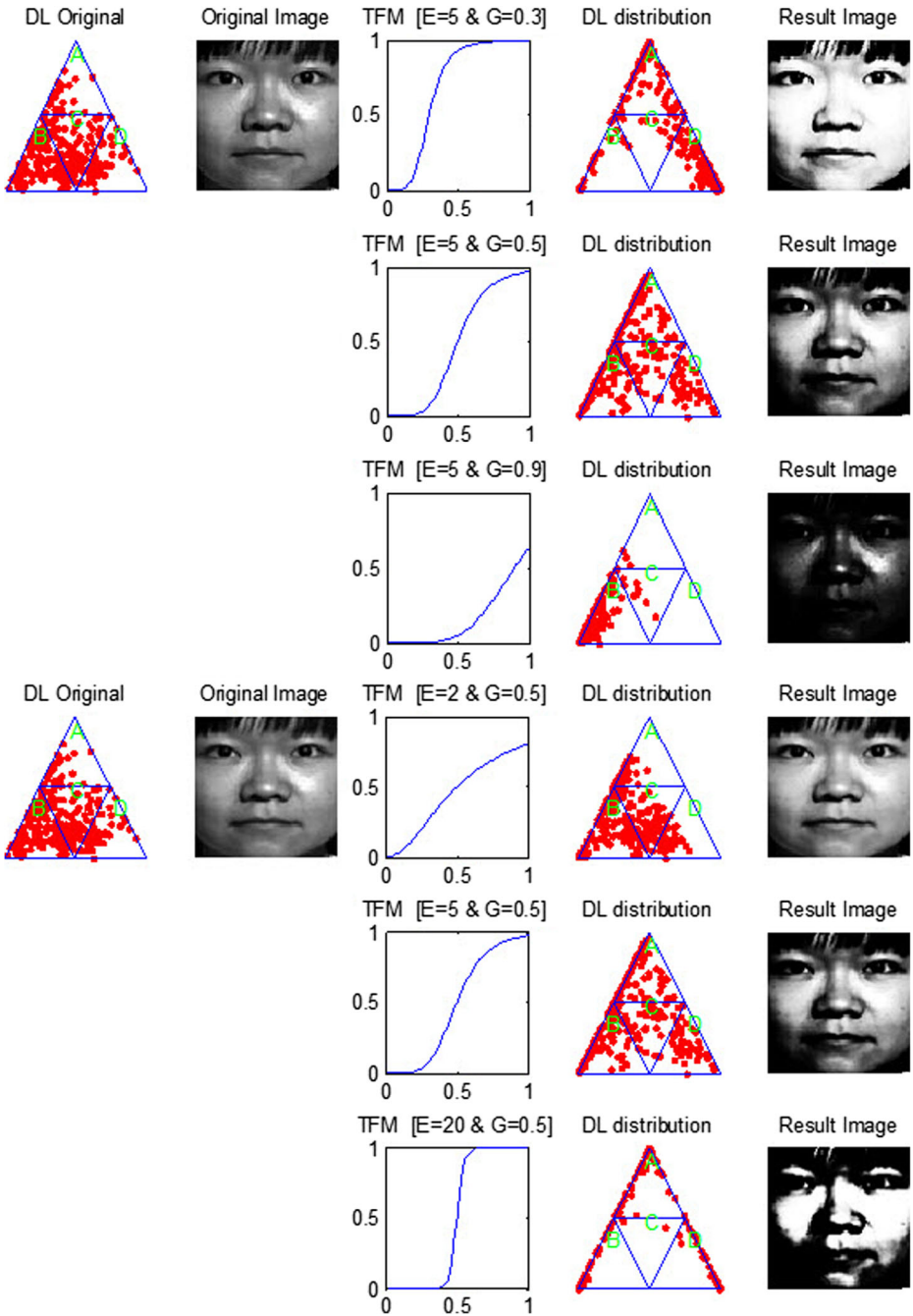


Fig. 5 TFM and DL descriptors distribution of CS function with variable factors values of the input/output image, transformation process with **a** $E = 5$ and $G = (0.3, 0.5 \text{ and } 0.9)$, **b** $G = 0.5$ and $E = (2, 5 \text{ and } 20)$

By variation of both G and E values, the pair (E, G) was inspected to understand its behavior and effect to the CS function via TFM and DL distribution analysis. By taking varying values of $G = 0.9, 0.5$ and 0.3 and a fixed value of $E = 5$ for the first case, various TFM curves were formed, which in turn affected the quality of the resulting image as shown in Fig. 5a. As compared to curves generated with a gain value of $G \leq 0.5$, curves generated with value $G > 0.5$ has a more distinct effect. The output face image appears to have a higher contrast than the original image and was observed to be darker when $G > 0.5$, following visual interpretation of the result images and TFM analysis. Conversely, when the value of $G \leq 0.5$, the output image is observed to be brighter and has lower contrast than the original face image. A fixed value of $G = 0.5$ with varying values of the factor $E = 2, 5$ and 20 are considered for the second case. As shown in Fig. 5b, the resulting filtered image is observed to have lower contrast and brighter than the original image as E increases. The filtered image is a binary image in a constraint limitation case with $E \geq 20$, the filtered image is a binary image. This limiting transformation is called the threshold transformation.

Prior to CS implementation, the DL scatter of original image is assumed to have the LC/HL description as shown in Fig. 5a, b. Observations made while the experiment for the value of $G = 0.3$ (case $G < 0.5$), noted that certain DL points in region D and C had concentrated and remained in region D and A. This resulted in the production of an image with maximum brightness and zero contrast. In contrast, the points moved in an opposite direction while points of regions D and C moved to region B when the value of $G = 0.9$ (case $G > 0.5$). This resulted in the production of an image with low brightness and contrast. Varying factors of $E = 2, 5$ and 20 is considered in the second case. As the E factor increases, the DL points in region D and C relocated to regions C and A. When $E \geq 20$, the points in regions C and A became concentrated and did not move away from its region.

Thus, the TFM and DL descriptors analysis and visual interpretation conclude that the pair (G, E) plays an important role in the brightness and contrast adjustment using CS. It also shows that the need of an appropriate algorithm to find the adaptive (G, E) values has a possible important function in the nonuniform illumination and contrast face enhancement via adaptive contrast, stretching along with giving an acceptable performance of an automated face recognition system..

2.2 Proposed approach overview

Figure 6 depicts the functional block diagram of the proposed ACS filter design via DL descriptors features and vocabulary tree model (VTM) in both training and testing phases. Archived DL descriptor features stored in the relational database are produced via off-line processing of face images in a collection. As shown in Fig. 6a, a training phase is applied to extract the DL features of face images, where the generated labels (G, E) in the database is achieved by Frame Differencing Module (FDM). A list DL descriptors feature associated with their corresponding labels generated from relevant face images is collected to build the VTM model. Additionally, the testing phase then performed to predict adaptively the face labels value. The results applied to the CS function to enhance the relevant face images as shown in Fig. 6b.

2.2.1 Frame differencing module (FDM)

The general idea behind the proposed method is to compute the pair factor (G, E) values from a Frame Differencing Module (FDM) approach in order to construct the vocabulary tree model (VTM). The FDM framework consists of two phases, namely i) E-generation

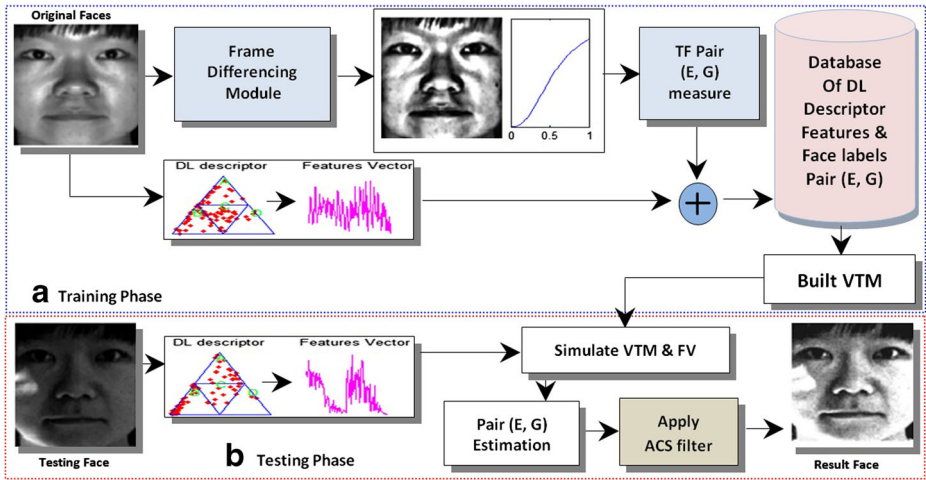


Fig. 6 Functional block diagram of the proposed ACS face enhancement

via transformed enhancer (E-TE) and (ii) G-generation via DL descriptors (G-DL). In the first phase, the image transformed enhancer (TE) involves the normalization of face input image (I). Two steps image enhancement process are performed. The first one is enhance contrast of I using histogram equalization (HE_I). The second one is obtained by enhance contrast of I using Contrast-limited Adaptive Histogram Equalization (CLAHE_I). Next, the (TE_{I_{nor}}) is applied to the normalized input image (I_{nor}) as in equation below:

$$TE_{I_{nor}} = |(CLAHE_{I_{nor}} + HE_{I_{nor}}) - I_{nor}| \tag{3}$$

Figure 7 shows a contrast enhancement results of sample face image and its corresponding histograms (cliffs) and cumulative histograms (CH) in blue graph-line using the FDM process.

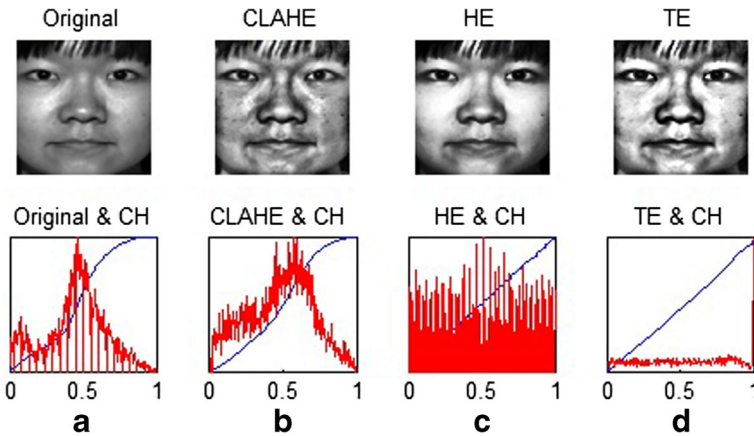


Fig. 7 Illustrate the **a** original image and **b, c, d** contrast enhancement using CLAHE, HE, TE, receptively with their corresponding histogram (cliffs) and cumulative histogram (blue graph-line)

Next, the steps involved in the E-TE process are as follows:

First step is to calculate the value of each histogram bin to find the histogram account (h_A) and bin location (b_L) using normalized transform of histogram on the resulting images $CLAHE_{I_{nor}}$, $HE_{I_{nor}}$, and $TE_{I_{nor}}$ with intensities values in the range of [0, 1].

The second step is to calculate the normalised cumulative sum ($CuSum_{Nor}$) of each resulting image histogram account (h_A) as follows:

$$CuSum_{Nor} = CuSum(h_A) / sum(CuSum(h_A)); \tag{4}$$

The third step is to find the coefficients of a polynomial $P(b_L)$ of degree $N = 1$ that fits the data ($X = CuSum_{Nor}$) best in a least-squares sense of each resulting image. P is a row vector of length $N + 1$ containing the polynomial coefficients in descending powers, $P(1) \times X^N + P(2) \times X^{(N-1)} + \dots + P(N) \times X + P(N + 1)$.

Then, the E factor is obtained using the following equation

$$E = P_{TE}(1) \times X^N + P_{CLAHE}(1) \times X^N + P_{HE}(1) \times X^N \tag{5}$$

In the second phase (G-DL), the G-factor is generated based on DL descriptors of the resulting $TE_{I_{nor}}$ and substituted in the CS function as in (1). Here, three different modes were examined to generate the G factor.

1) The first mode (G_{M1}) considered as the mean (μ) of the DL descriptors values extracted from the $TE_{I_{nor}}$ given as:

$$G_{M1} = \mu = \frac{\sum_{i=1}^k f_i(x_i)}{n - 1} \tag{6}$$

2) The second mode (G_{M2}) considered as the stand deviation (σ) of DL descriptors values extracted from the $TE_{I_{nor}}$ given as:

$$G_{M2} = \sigma = \sqrt{\frac{\sum_{i=1}^k \frac{f_i(x_i - \mu)^2}{n - 1}}{n - 1}} \tag{7}$$

3) The third mode (G_{M3}) considered as the variance (σ^2) of DL descriptors values extracted from the $TE_{I_{nor}}$ given as:

$$G_{M3} = \sigma^2 = \sum_{i=1}^k \frac{f_i(x_i - \mu)^2}{n - 1} \tag{8}$$

where, $f_i(x_i)$ is DL descriptors set values generated from $TE_{I_{nor}}$, and n is the number of set values.

2.2.2 Vocabulary tree model (VTM)

Using a hierarchical k-means cluster that categorises the connected classes from the test data quickly with good performance, the VTM forms a hierarchical quantization threshold (cluster centre) [26]. At a training step, the most commonly adopted unsupervised learning method is constructing a tree by utilising a K-means clustering. Subsequently, the tree is expanded- where the accuracy rate KLD, in which K, L and D represent the branching factor, depth of the tree and extracted input vector dimension respectively. Regardless of whether the data sets have a class label or not, the VTM can still be utilised. Here,

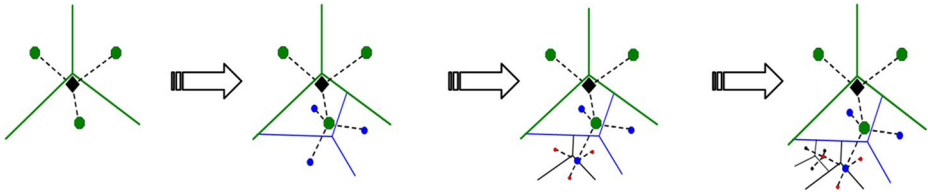


Fig. 8 Process flow to construct the VTM, where $K = 3$ define or specify the centre of the hierarchical quantization for each level [26]

the DL descriptors vectors extracted from intensity histogram of TE_{Inor} associated with E factor that is generated via E-TE process are utilised to compute the VTM. The steps of constructing vocabulary tree are shown in Fig. 8.

The building and testing of VTM involves the following three major steps; (1) construction of tree model based on k-means clustering, (2) mapping of the DL descriptors vectors on the tree and (3) calculating the score of the action recognition.

2.3 Vocabulary tree construction

For this module, a DL feature vector and its pair set (G, E) class labels represent each training vector. K-means clustering technique is used to partition these labelled training sets into different groups. A cluster refers to the pool of the data object within the exact same cluster possesses similar features but dissimilar to the objects in other clusters. Many parameters that strongly influence the classification results, which includes parameters such as the depth of the tree L and branching factor K , affects the VTM construction [26]. Thus, the value of L is fixed as $L = 4$ while the value of K is examined with different values as $K = 3, 5, 7$ and 9 , to evaluate the VTM decision system. DL descriptor features are used to extract all feature vectors from the training data sets. The VTM construction as detailed out in Algorithm 1 is illustrated in Fig. 9.

Algorithm 1 VTM construction

- 1-The constant K and the functions that categorise a K -children node are determined..
 - 2-The tree depth level $L = 4$ is defined.
 - 3- M is defined as the number of DL vector points selected as training vectors.
 - 4-DL vector points (M) are arranged into k clusters making use of k-means clustering approach throughout level $L = 1$ clustering.
 - 5-At each leaf node, the respected cluster centre is stored along with their cluster members.
 - 6-Identical process will be repeated recursively to suit each cluster relative to the vector points, hence the cluster members of the respective groups tend to be more partitioned into k -groups.
 - 7-This recursive practice continues until the maximum $L = 4$ levels are reached.
-

2.4 DL feature vector mapping

Every DL vector points of the face images are mapped on the VTM [26], where will be different cluster centre of each group for all different levels. Figure 10 shows the DLs feature vector mapping process as detailed out in Algorithm 2.

Algorithm 2 DL feature vector mapping

- 1- Each (M) point is distributed down the tree by comparing each component with the k -cluster centres for each level. The selection is based on the closest distance using L2 metric.
- 2- Identical process is replicated for all DL vector points in the face image.
- 3- Steps (1) and (2) are replicated for the whole image database.
- 4- Feature vectors numbers (n^{th}) within the leaf node (i^{th}) are counted and saved. This step is repeated for all leaf nodes.

2.4.1 Score and action recognition

The label E is then recognised by utilising the mapping. Based on different level (L^{th}) and (i^{th}) leaf nodes, each face image DL feature description will be defined. At this particular stage, the DL feature vector of the tested image (F_T) and the DLs feature vectors of the entire database (F_I) are defined according to the following equations:

$$F_T = n_i w_i \tag{9}$$

$$F_I = m_i w_i \tag{10}$$

$$w_i = \ln \frac{N}{N_i} \tag{11}$$

where, m_i and n_i are feature description numbers of the query image and the database DLs passing via leaf node (i), w_i represents the weight of each node i , N is the total number of database DLs and N_i is the the total number database DLs in conjunction with at least one vector point passing via the node i . Similarity normalization is employed to ensure fairness among database DLs having different vector points and given as:

$$F_T = \frac{F_T}{\|F_T\|} \tag{12}$$

$$F_I = \frac{F_I}{\|F_I\|} \tag{13}$$

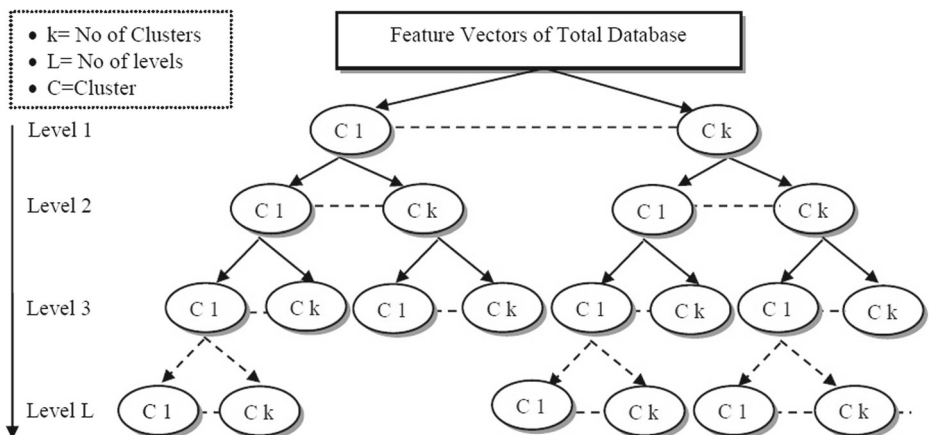


Fig. 9 Vocabulary tree model (VTM) construction

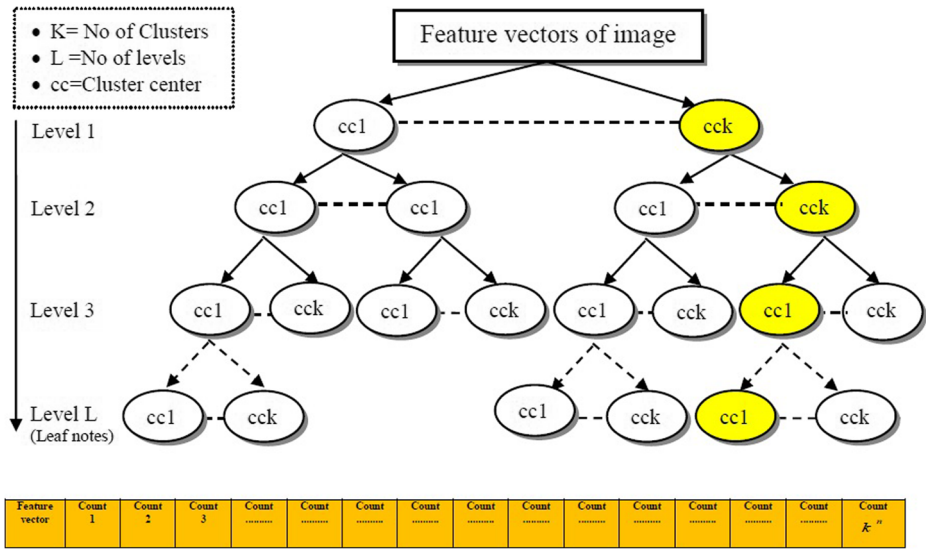


Fig. 10 DL feature vector mapping on the VTM

At the action recognition stage, an face label index E is utilized to represent similarity measure which is computed using the testing feature T_i along with the database (I_i) given by:

$$E = \sum_{i=1}^{L=4} F_T \times F_I \tag{14}$$

With regards to stored features point from the database DLs feature vectors, the resulting feature points from the query DLs feature are assessed. Should the maximum feature point correspond to the query feature point, the relevant features are confirmed as the correct feature. The relevant E-factor is taken as the class label of the matched query feature point that is obtained. Algorithm 3 gives the implementation processes of the concept used to get the score value and recognise the action.

Algorithm 3 Score and action recognition

- 1-Extract the features vector (DLs) from the tested face image.
- 2-Generate the class label of the corresponding DLs using equation (12).
- 3-Label the queried feature represented by a pair set [DL, E].
- 4-Load the resulting VTM which models the labelled feature vectors from the database.
- 5-Compute the score E by comparing the query feature with the database feature vectors.
- 6-Sort the scores, where the highest feature point that matched the query feature point is the best.
- 7-Obtain the class label of the matched query feature point.
- 7-The resulting class label will be considered as the relevant E-factor.

2.5 VTM performance evaluation

The VTM decision system is evaluated using the KLD index investigation, in order to assess the performance of the suggested algorithm. This investigation takes the depth of the tree, $L = 4$ with a set of branching factor values $K = 3, 5, 7$ and 9 with DL feature vector dimension $D = 162$ into consideration. The area under the curve (AUC) measurement taken during the VTM testing stage is utilised to generate the receiver operating characteristic (ROC) and VTM training model to give quantitative evaluation for the selected KLD configuration. A systematic analysis of specificity and sensitivity of the diagnosis is provided by the ROC graph [6, 37].

The true positive and true negative indexes are represented by TP and TN, respectively that represent the agreement between the classifications of the professionals as compared to the VTM classifier. In contrast, the false positive and false negative indexes are represented by FP and FN, respectively that represents the disagreement between VTM classification and the professionals. When all approval patterns are introduced on the VTM classifier at the end of every epoch, the statistical record indices of that epoch (e) will be measured for each threshold t as well as sensibility and specificity [40]. Equation (15) gives the sensitivity (S_n), that is plotted along the abscissa axis to signify the VTM classifier abilities to identify the positive pattern among the truly positive patterns when S_n varies from 0 (TP = 0, FN ≠ 0) to 1 (FN = 0, TP ≠ 0), a higher test sensitivity is represented by the smaller number of false negatives.

$$S_n(e, t) = \frac{\sum TP(e, t)}{\sum TP(e, t) + \sum FN(e, t)} \quad (15)$$

Equation (16) gives the specificity (S_p) that is plotted along the ordinate axis to signify the ability of the VTM classifier to identify the negative pattern among the truly negative patterns. S_p varies between [0, 1] where a higher test specificity is represented by a smaller number of false negative.

$$S_p(e, t) = \frac{\sum TN(e, t)}{\sum TN(e, t) + \sum FP(e, t)} \quad (16)$$

AUC as in (17) represents a significant measure of the accuracy of the VTM test. AUC equals to 1 for the ideal classifier process $S_n = 1$ (TP = 1) and $S_p = 1$ (FP = 0) [6]. Thus, should both S_n and S_p are 1, the test is considered 100% accurate. AUC falls within the range of $0 \leq AUC \leq 1$, in which $AUC = 1$ represents a perfect test accuracy while $AUC = 0.5$ reflects a random test accuracy. As such, it is likely that $AUC < 0.5$ signifies a rejection in model test.

$$AUC = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n 1_{p_i > p_j} \quad (17)$$

where, all n data points with true label 0 (predict incorrect) are covered by j ; all m data points with true label 1 (predict correct) are covered by i , all n data points with true label 0 (predict incorrect) are covered by j ; p_i and p_j represents the probability score assigned by the VTM classifier with relation to the data point i and j , respectively. 1 represents the indicator function satisfied by the condition ($p_i > p_j$).

2.6 Image quality measurement (IQM)

According to the availability of a reference image, objective evaluation techniques are classified as Full-Reference (FR), Blind-Reference (BR) and Reduced-Reference (RR) image quality metrics [22]. In this paper, the proposed ACS is examined out on four publicly available databases and the face quality evaluation was measured via BR based statistical measures (BR-SM), BR based DL descriptors (BR-DL) and visual interpretation of the resulting images.

2.6.1 IQM via BR-SM

Quantitatively, six quantitative evaluations, which are EME (measure of enhancement), EMEE (measure of enhancement by entropy), SDME (Second Derivative like Measure of Enhancement), Sharpness (the mean of intensity differences between adjacent pixels taken in both the vertical and horizontal directions), and CPP (Contrast Per Pixel) are selected to objectively evaluate the proposed ACS relative to other contemporary methods.

The EME and EMEE [3, 4] given by (18) and (19), respectively has been worked upon by Agaian et al., following several modifications of the Webers and Fishers Laws. EME provides an absolute value to each image by evaluating image contrast using Webers law and linking it to the perceived brightness based on Fishers law.

$$EME = x \left(\frac{1}{k_1 k_2} \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} 20 \ln \left[\frac{I_{max;k,l}^w}{I_{min;k,l}^w + c} \right] \right) \tag{18}$$

$$EMEE = x \left(\frac{1}{k_1 k_2} \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \alpha \left[\frac{I_{max;k,l}^w}{I_{min;k,l}^w + c} \right]^\alpha \ln \left[\frac{I_{max;k,l}^w}{I_{min;k,l}^w + c} \right] \right) \tag{19}$$

where, image I is divided in to $(k_1 \times k_2)$ blocks $wk, l(i, j)$ of size $(l_1 \times l_2)$, I_{min} and I_{max} are the maximum and minimum values of the pixels in each block. Generally $c = 0.0001$ and α is between 0 to 1. Function x is sign function (– or +) depending on the used enhancement method [3].

The concept of the second derivative is utilised by SDME [46] in addition of measures (maximum and minimum pixel values) take into account, it also considers centre pixel value. Compared to other measures, it is less sensitive to noise and steep edges [46].

$$SDME = - \frac{1}{k_1 k_2} \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} 20 \ln \left[\frac{I_{max;k,l} + 2I_{center;k,l} + I_{min;k,l}}{I_{max;k,l} + 2I_{center;k,l} + I_{min;k,l}} \right] \tag{20}$$

In addition, the sharpness *SHP* assesses the filter’s capacity to steadily decrease the noise as defined by Kryszczuk and Drygajlo [18] and given as:

$$SHP = \frac{1}{2} \left[\frac{1}{(N-1)M} \sum_{i=1}^M \sum_{j=1}^{N-1} |I_{i,j} - I_{i,j+1}| + \frac{1}{(M-1)N} \sum_{i=1}^{M-1} \sum_{j=1}^N |I_{i,j} - I_{i+1,j}| \right] \tag{21}$$

where, $I(i, j)$ indicates the initial image represented by $M \times N$ dimensions.

Contrast Per Pixel (CPP) [7, 10] refers to an estimation of the difference in average intensity between a pixel and its adjacent pixel. The following equation defines the CPP of an image:

$$CPP = \frac{\sum_{i=0}^M \sum_{j=0}^N (\sum_{(m,n) \in R_3(i,j)} |I(i,j) - I(m,n)|)}{MN} \quad (22)$$

where, $I(i, j)$ is the gray value of pixel (i,j) , and $I(m, n)$ is the gray value of neighboring pixel (m, n) in the 3×3 window.

These parameters show that the smoothing capability of the ACS filter, having a higher (EME, EMEE, SDME, SHP and CPP) implies a better image quality.

2.6.2 IQM via BR-DL

The visual impression may not correspond accurately with the evaluation results of BR-SM set measure. Thus, an extra assessment approach is utilised to assess different perspective of the filter performance, especially for the component that manages the partnership between the local contrast, LC and local brightness, LB using location and dispersion estimators [24, 32, 35]. The BR-DL assessment based on the range and midrange pairs computation from the four data sets before and after applying the seven contrast enhancement methods. First, generated the points in the scatter plot that fall in each of the four region triangles A, B, C, and D, which subdivide the basic triangle in the dispersion-location plane. Next, compute the mean of the points in percentage $\mu(\%)$ corresponding to each region before and after employing the different contrast enhancement methods. This static $\mu(\%)$ distribution in each region could help explains not only the changes in nonuniform luminance that the images undergo but also the changes in contrast with the relationships changes between the contrast and luminance in the statistical measurement indicated by specific description. Thus, the $\mu(\%)$ of the DL scatter plot distribution at each region is measured to evaluate the resulting enhanced face images. Alternatively, the face images are evaluated qualitatively by finding and visual observing the directional gradients of the resulting enhanced image clarity.

3 Results and discussion

3.1 Optimisation of G factor

In the face images from the EYale-B, the CPP values are plotted as shown in Fig. 11a, and Table 2 highlights the averaged CPP values. Compared to those in the original image, the three G factor processed images have an increased CPP values. This means the ACS processing by all the modes resulted in enhanced contrast. The CPP of the image EYale-B processed by ACS with Mode 3 is highest and over-enhanced contrast. This phenomenon is due to over-stretching of the histogram and causes the loss of the detailed information, especially in the dark or bright tissue contents. The CPP values of the images processed by ACS with Mode 1 and Mode 2 are modest. This can be proven by images comparison, which suggests that the contrast is well enhanced for images both poorly-stained and well-stained, both dissimilar and similar when compared to the reference image, this results in an imprecise grey value transformation, losing some grey levels in the images differing from the original image.

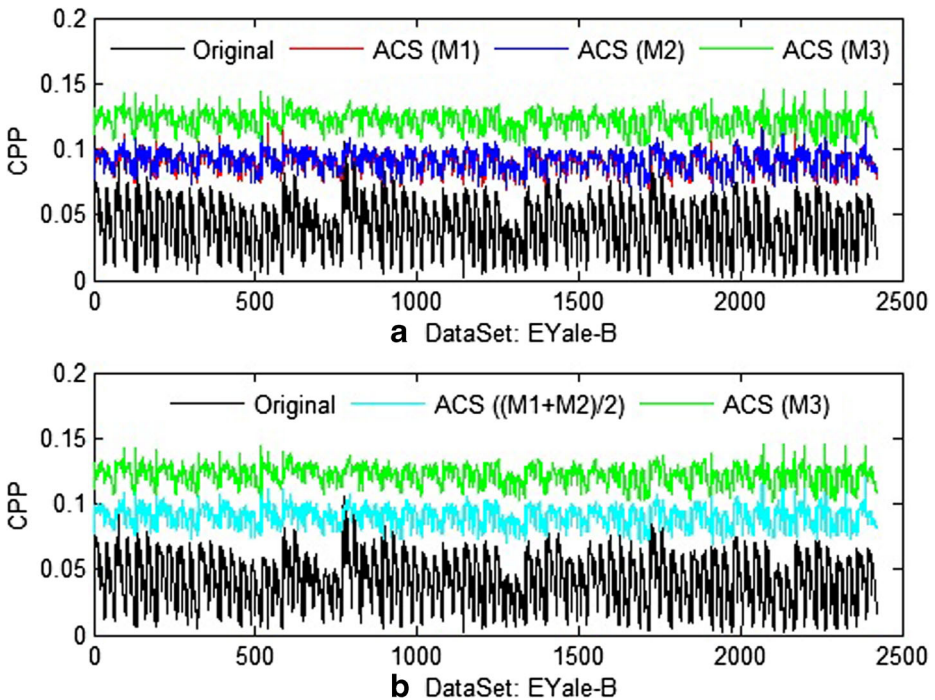


Fig. 11 CPP of the original and ACS enhanced images with three G modes using EYale-B database

By averaging both ACS with Mode 1 and Mode 2 as shown in Fig. 11b, the global contrast of processed images is the most enhanced among all EYale-B. The contrast of the background is well-enhanced and processed image have increased compared to those in the original image when dealing with such images by averaging mode. This implies that the contrast has been well enhanced following the processing by average mode.

Figure 12 depicts the face images with corresponding histograms (cliffs) and cumulative histograms (blue graph-line) using the adaptive contrast stretching with different G factor mode measure.

3.2 ROC and VTM performance

In this section, the VTM ability to model the DL descriptors features for E factor prediction was examined and evaluated by employing the ROC graph and AUC scores index. The proposed ACS filter was tested on four data sets (EYale-B, Mobio, Feret and CMU-PIE), where the patterns of each data set image are separated into two sets. The training set is two third of the images from each data set where the remaining face images were used to evaluate the VTM. In the training phase, the normalised DL descriptors are extracted from each image intensity histogram as the input to the VTM, where their corresponding E-factors are generated via FDM. Increasing branching factor K has also been used to increase the nodes in the vocabulary tree. Therefore, several experiments were examined to obtain the best K value by increasing branching factor $K = 3, 5, 7$ and 9 with the depth of $L = 4$. Figure 13 show the ROC graphs and AUC scores of VTM performance with different K values.

Table 2 Image quality measurement (IQM) via BR-SM

Databases	Original	ACS	CLAHE	HE	BPDFHE	NMHE	WA	BEASF
EYale-B								
EME	19.3031	29.6762*	26.0194	18.8186	20.8964	22.5206	28.1083	22.5218
EMEE	34.9230	52.2587*	44.3383	34.6036	35.7708	46.1310	39.2518	48.5794
SDME	0.1043	0.1858*	0.1385	0.1291	0.1029	0.1278	0.1598	0.1295
SHP	0.0262	0.0584	0.0502	0.0422	0.0286	0.0416	0.0599*	0.0415
CPP	0.0440	0.0902*	0.0542	0.0744	0.0467	0.0737	0.0738	0.0737
Mobio								
EME	15.3504	31.8897*	23.1459	19.5925	16.0919	23.8638	30.7188	23.8057
EMEE	38.1234	52.1946*	38.1475	39.1143	38.9769	44.7564	36.1724	45.4420
SDME	0.0862	0.2020*	0.1184	0.1324	0.0876	0.1305	0.1436	0.1306
SHP	0.0257	0.0581*	0.0506	0.0410	0.0274	0.0406	0.0524	0.0406
CPP	0.0582	0.1008*	0.0704	0.0778	0.0599	0.0777	0.0776	0.0777
Feret								
EME	14.5711	34.5963*	23.0260	21.6677	15.6728	25.8408	34.2521	25.8144
EMEE	40.2820	52.5664*	38.4539	40.5262	40.7583	46.6260	39.2888	46.6579
SDME	0.1003	0.2255*	0.1343	0.1553	0.1010	0.1529	0.1679	0.1529
SHP	0.0272	0.0675*	0.0549	0.0491	0.0297	0.0485	0.0596	0.0485
CPP	0.0604	0.0954*	0.0712	0.0727	0.0620	0.0727	0.0727	0.0727
CMU-PIE								
EME	20.5210	40.4475*	29.0190	21.4605	22.1526	31.4684	32.3545	31.6251
EMEE	40.0820	57.3548	53.8819	36.1401	36.9958	57.6117	55.8255	58.8863*
SDME	0.1309	0.2417*	0.1700	0.2205	0.1309	0.2119	0.2288	0.2128
SHP	0.0272	0.0724*	0.0517	0.0624	0.0304	0.0571	0.0702	0.0574
CPP	0.0288	0.0832*	0.0449	0.0722	0.0321	0.0723	0.0725	0.0722

The displayed results in Fig. 13 show that VTM decision system with DLs descriptors vector extraction of all data sets delivered satisfactory E-factor prediction with $K=7$ for the lower test accuracy of 0.8652; while the higher AUC test accuracy of 0.8739 is obtained at $K=5$. Generally, all the results obtained show that the E-factor prediction performed by the VTM decision system to recognise the DLs feature vectors have displayed satisfactory performance with average AUC score of more than 0.865.

3.3 IQM results interpretation

As indicated earlier, the ACS filter performance and face quality measurement against the state-of-the-art methods was examined via BR-SM, BR-DL matrices and visual interpretation of the resulting images.

3.3.1 Assessment via BR-SM

The IQM via BR-SM has been computed based on the predicted pair (E, G) substituted in the CS function. Figures 14, 15, 16, 17 and 18 produced with receptively, EYale-B, Mobio, Feret, and CMU-PIE data sets and highlights the performance of the suggested approach

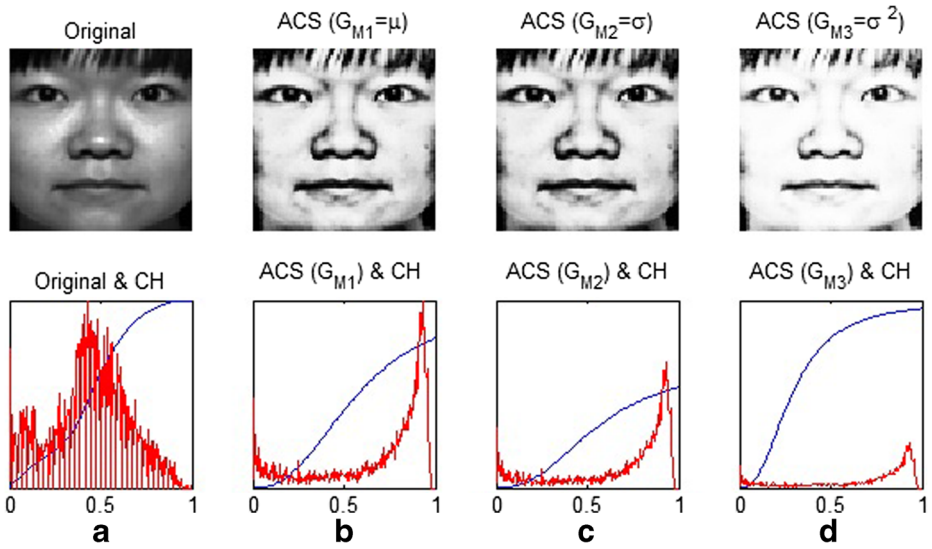


Fig. 12 ACS Corresponding histogram (cliffs) and cumulative histogram (blue graph-line) of **a** face original sample and ACS results with **b** $G_{M1} = \mu$, **c** $G_{M1} = \sigma$ and **d** $G_{M1} = \sigma^2$

ACS and the six other contemporary enhancement methods (CLAHE, HE, BPDFHE, NMHE, WA, and BEASF) quantified by BR-SM quality metrics (EME, EMEE, SDME, SHP, and CPP). In order to make comparison, the different contrast enhancement methods are arranged from left to right in an increasing performance order, while the original data sets and

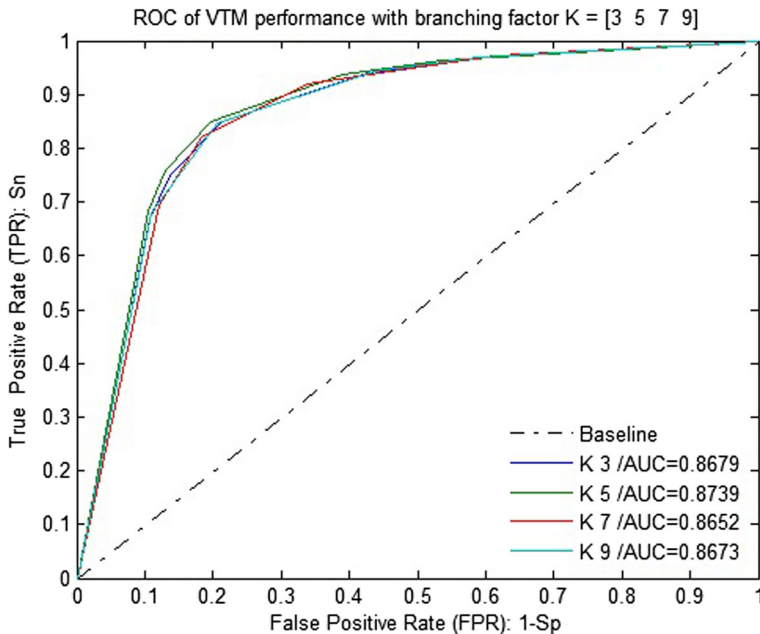


Fig. 13 ROC of VTM performance with branching factor $K = 3, 5, 7, 9$

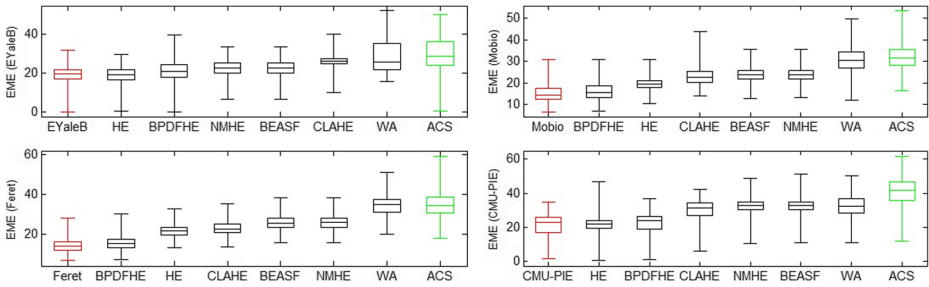


Fig. 14 Mean value of EME metric for four original databases and each of the seven enhancement methods (ACS, CLAHE, HE, BPDFHE, NMHE, WA, and BEASF) assessed in this study. The proposed methods ACS and databases are represented with *green and red colors symbol*, respectively to enable easy comparison. The standard error of the mean is represented by the *error bars*

proposed ACS are represented with red and green color symbols, respectively. Figures 14 and 15, shows receptively the mean value of EME and EMEE metrics with the overall data sets using the ACS along with the six contrast enhancement methods.

These results based EME and EMEE measures highlights that the ACS and WA approach obtained the largest overall mean rankings respectively, which implies that the four data sets processed by these methods were objectively rated as looking the most natural and possessing a satisfactory overall enhancement. CLAHE, NMHE and BEASF obtained intermediate ranking values, while BPDFHE and HE obtained the lowest rankings. The relatively high rating of ACS is due to the participants appreciating the large enhancement of the dark regions in several images of the test data sets. The proposed ACS achieved the best results in EME and EMEE for all types of data sets, which plainly demonstrated that ACS has higher contrast enhancement capability as compared to other contemporary methods. Figure 16 highlights the resulting SDME mean value of the similar data sets using the seven contrast enhancement approaches assessed in this study.

Results in terms of SDME metric, method BPDFHE seems to be the most under-performing techniques in all data sets, while the ACS seems to be high within the performance range. The objective mean ranking of ACS and WA, induced by the SDME metric are in line to a large extent with the objective ranking order by the EME and EMEE metrics indicated above. Note that, an increase in the face sharpness correlates with an increase in the observed SDME.

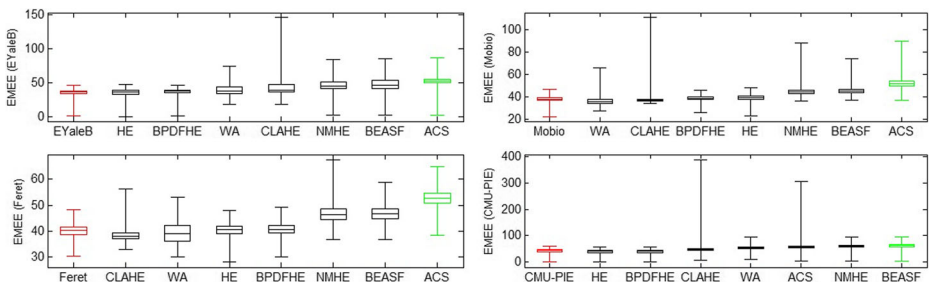


Fig. 15 Mean value of EMEE metric for four original data sets and each of the seven enhancement techniques (ACS, CLAHE, HE, BPDFHE, NMHE, WA, and BEASF) investigated in this study. The proposed method ACS and data sets are represented with *green and red colors symbol*, respectively to allow easy comparison. The standard error of the mean is represented by the *error bars*

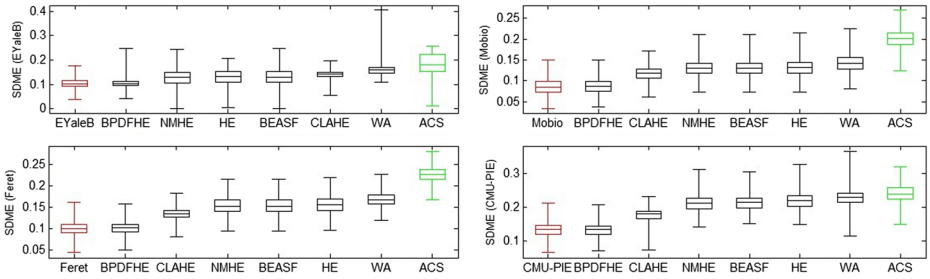


Fig. 16 Mean value of SDME metric for four original data sets and each of the seven enhancement techniques (ACS, CLAHE, HE, BPDFHE, NMHE, WA, and BEASF) assessed in this study. In order to allow easy comparison. The proposed method ACS and data sets are represented with green and red colors symbol, respectively. The standard error of the mean is represented by the error bars

In addition, shown in Fig. 17 is the mean value of sharpness (SHP) metric measured from the four data sets using the seven contrast enhancement techniques. The SHP results demonstrate that ACS and WA methods have an overall mean SHP values higher than other methods, proving that both approaches assessed do enhance image contrast effectively. ACS causes the most enhancement in image contrast and edge sharpness as compared to WA, since it has the largest mean SHP values. Again, it should be noted that the objective order induced by the SHP metric is in line to a large degree with the visual inspection of sample images for different subsets in the four data sets.

Furthermore, Fig. 18 shows the mean CPP values with the four data sets using the seven contrast enhancement techniques assessed in this study. ACS and WA induce the largest improvement in image contrast and edge sharpness since these methods show the largest mean CPP values. It is likely that the relatively high rating of ACS is due to the participants appreciating the significant improvement of the dark regions in several images of the data sets.

In summary, these results based BR-SM highlights that the proposed ACS enhance face structure effectively while maintaining the features of the original image to a large extent. ACS is also comparable to a few contemporary contrast enhancement methods with respect to contrast enhancement and the preservation of face structure. Summarised detail of measurement results are given in Table 2, where the IQM via BR-SM of the ACS are proven to have relatively almost the high mean value of EME, EMEE, SDME, SHP and CPP metrics

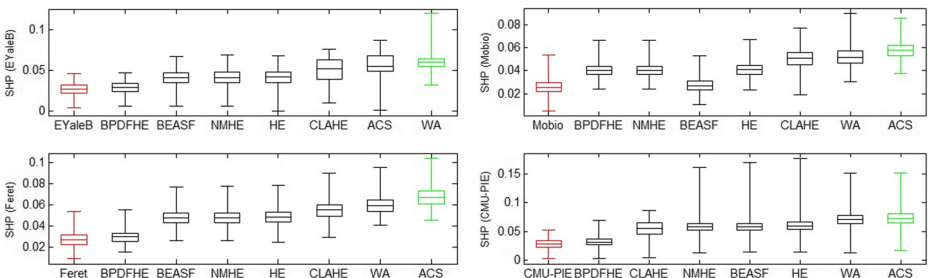


Fig. 17 Mean value of SHP metric for four original data sets and each of the seven enhancement techniques (ACS, CLAHE, HE, BPDFHE, NMHE, WA, and BEASF) assessed in this study. In order to allow easy comparison. the proposed method ACS and data sets are represented with green and red colors symbol, respectively. The standard error of the mean is represented by the error bars

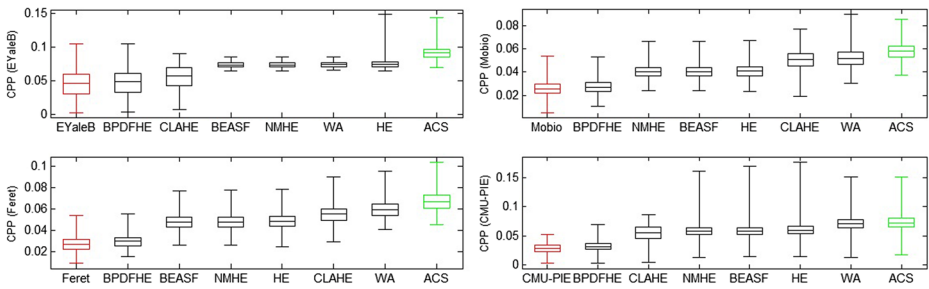


Fig. 18 Mean value of CPP metric for four original data sets and each of the seven enhancement techniques (ACS, CLAHE, HE, BPDFHE, NMHE, WA, and BEASF) assessed in this study. In order to allow easy comparison, the proposed method ACS and data sets are represented with green and red colors symbol, respectively. The standard error of the mean is represented by the error bars

compared to the six contemporary enhancement methods which are found to have the almost the lowest measurements.

3.3.2 Assessment via BR-DL

The IQM via BR-DL has been also computed based on the predicted pair (E, G) substituted in the CS function. Figure 19 illustrates the $\mu(\%)$ DL distributions at each region from the original EYale-B data set before and after applying the proposed ACS along with the six contrast enhancement methods investigated in this study.

As shown in Fig. 19a, the DL distributions presented by $\mu(\%)$ in each region from the original EYale-B images before enhancement, with distributions descriptors BCDA occurring in decreasing order. The underlined letter indicate the presence of $\mu(\%)$ points in region B was higher compared with that in other regions (the points in region B were concentrated more than the points in regions C, D, and A had lower points compared with region B), this signifies that the entire EYale-B images underwent dark with poor quality and generally considered as Low-Contrast/Low-Luminance (LC/LL) before enhancement.

By employing the proposed ACS method, significant distributions descriptors DABC (underlined letters indicate higher presence of $\mu(\%)$ points in both regions D and A) occurred for EYale-B data set, as shown in Fig. 19b. Implies, that points in regions D and A has been expanded, and the points in regions B and C has been decreased, unlike the movements of the points when using the original distributions descriptor BCDA. These movements indicate that the entire EYale-B data set enhanced with the luminance amount adjusted and contrast increased. Hence, the entire EYale-B data set generally enhanced and considered as High-Contrast/Medium Luminance (HC/ML) after employing the ACS.

Figure 19c, d, e, f, g, h respectively, shows the distributions descriptors of the six contrast enhancement methods (CLAHE, HE, BPDFHE, NMHE, WA and BEASF) with the DL distributions descriptors receptively, BACD, BDAC, BCDA, BDCA, BDAC, and BDAC. Generally speaking, those $\mu(\%)$ had the new distributions descriptors occurred and indicated that almost the major points concentrated in region B, signified that the entire EYale-B data set enhanced slightly with medium to lower luminance (most cases, a higher presence of points in region B, and D) and had lower contrast (higher or total absence of points in regions C and A) compared with the distributions descriptors by employing the proposed ACS.

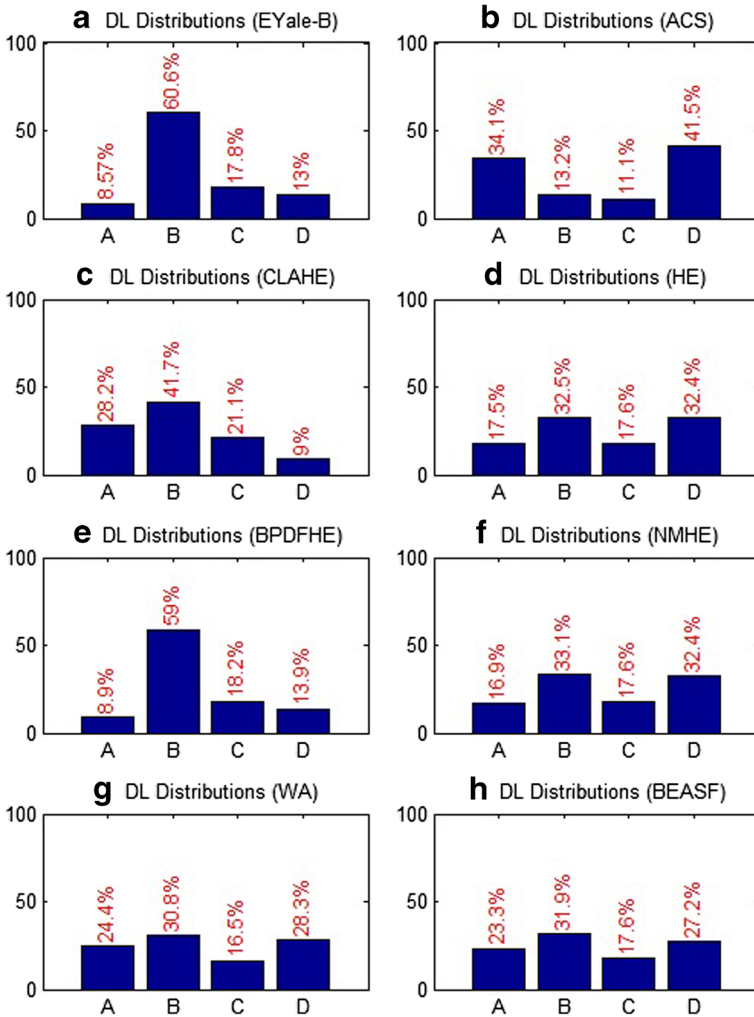


Fig. 19 $\mu(\%)$ of DL distributions at each region of the **a** original EYale-B data set before and after applying the seven contrast enhancement methods **b** ACS, **c** CLAHE, **d** HE, **e** BPDFHE, **f** NMHE, **g** WA, and **h** BEASF

Figure 20 illustrates the $\mu(\%)$ DL distributions at each region of the original Mobio data set, before and after applying the proposed ACS along with the six contrast enhancement methods.

The $\mu(\%)$ DL distributions from the original Mobio data set before enhancement with distributions descriptors *BCDA* occurring in decreasing order as shown in Fig. 20a. The description indicated that the major points concentrated in region B and signified that the entire data set undergoes dark with poor quality before enhancement. Thus, the Mobio data set is generally considered as low-contrast/Low luminance (LC/LL).

A significant distributions descriptors *DABC* occurred for Mobio data set by employing the proposed ACS as indicated in Fig. 20b. The points in regions D and A expanded, and the

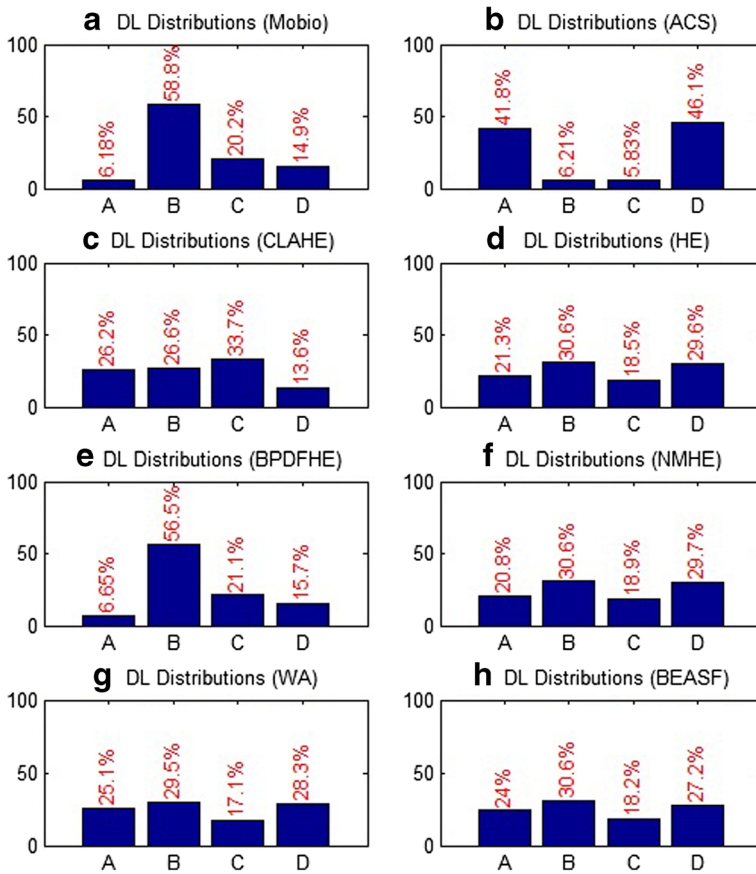


Fig. 20 $\mu(\%)$ of DL distributions at each region of the **a** original Mobio data set before and after applying the seven contrast enhancement methods **b** ACS, **c** CLAHE, **d** HE, **e** BPDFHE, **f** NMHE, **g** WA, and **h** BEASF

points in region B and C decreased, unlike the movements of the points when using the original distributions descriptor \underline{BCDA} . Here, the entire Mobio data set has been enhanced with luminance amount adjusted and contrast increased (all the $\mu(\%)$ points in D and A regions were higher compared to other regions). By employing the ACS, the entire data set has been enhanced and generally considered as High-Contrast/Medium-Luminance (HC/ML). Again, shown respectively in Fig. 20 (c, d, e, f, g, h), the distributions descriptors of the six contrast enhancement methods (CLAHE, HE, BPDFHE, NMHE, WA and BEASF) with the DL distributions descriptors receptively, \underline{CBAD} , \underline{BDAC} , \underline{BCDA} , \underline{BDAC} , \underline{BDAC} , and \underline{BDAC} . Generally speaking, the $\mu(\%)$ have a new distributions descriptors occurred and almost the major points concentrated in either region B or C (the $\mu(\%)$ in region B or C were higher compared with that in other regions). This signified that the entire Mobio data set enhanced slightly with medium to lower luminance (a higher presence of points in region B) and had lower contrast (higher or total absence of points in regions C and A) compared with the distributions descriptors using the proposed ACS.

Figure 21 illustrates the $\mu(\%)$ DL distributions of the original Feret data set before and after applying the proposed ACS along with the six contrast enhancement methods.

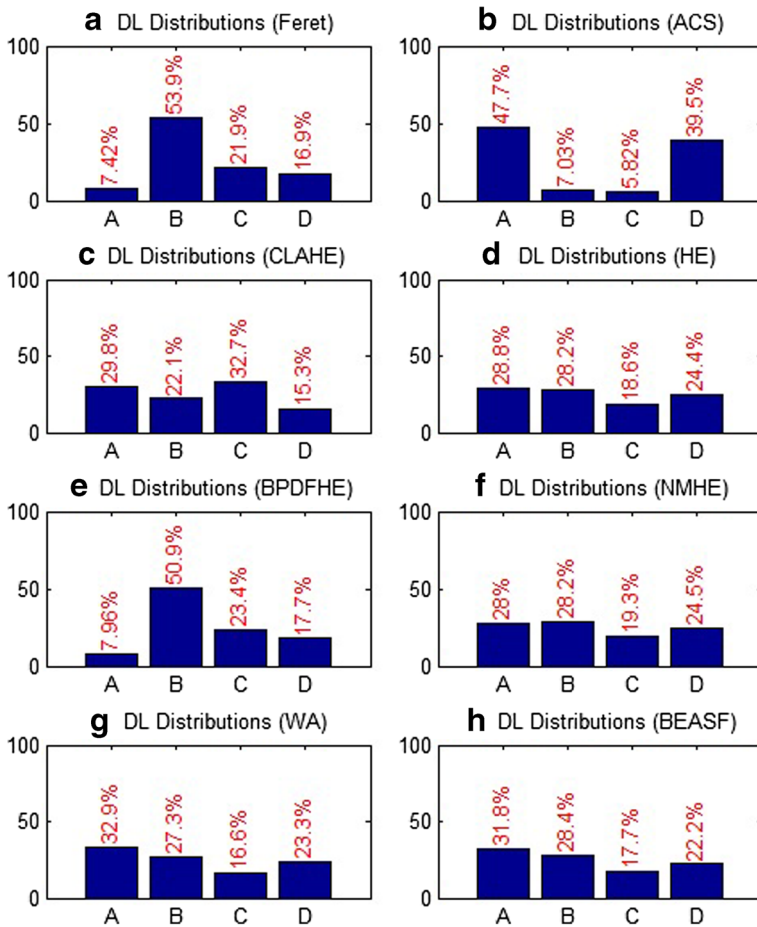


Fig. 21 $\mu(\%)$ of DL distributions at each region of the **a** original Feret data set before and after applying the seven contrast enhancement methods **b** ACS, **c** CLAHE, **d** HE, **e** BPDFHE, **f** NMHE, **g** WA, and **h** BEASF

Again, the $\mu(\%)$ DL distributions from the original Feret images before enhancement has distributions descriptors \underline{BCDA} occurring in decreasing order as shown in Fig. 21a, the $\mu(\%)$ in region B was higher compared with that in other regions, signified that the entire Feret dataset underwent dark and had poor quality. Hence, the entire data set was generally considered as Low-Contrast/Low-Luminance (LC/LL).

By employing the proposed ACS, a significant distributions descriptors \underline{ADBC} occurred for Feret data set as illustrated in Figure shown in Fig. 21b. Thus, the points in both A and D regions expanded and the points in region B and C decreased, unlike the movements of the points when using the original distributions descriptors. Indicated, that the entire data set enhanced with the luminance amount adjusted and contrast increased (all the $\mu(\%)$ points in A and D regions were higher compared to other regions). Hence, the entire Feret data set has been enhanced using the proposed ACS and considered as High-Contrast/Medium-Luminance (HC/ML).

Indicated in Fig. 21c, d, e, f, g, h are respectively the distributions descriptors of the six contrast enhancement methods (CLAHE, HE, BPDFHE, NMHE, WA and BEASF)

with DL distributions descriptors receptively, \underline{CABD} , \underline{ABDC} , \underline{BCDA} , \underline{BADC} , \underline{ABDC} , and \underline{ABDC} . Observed, those $\mu(\%)$ had a new distributions descriptors occurred and indicated that almost the $\mu(\%)$ in regions C, B and A were higher compared with that in other regions), signified that the entire Feret data set enhanced slightly with medium to lower luminance (most cases, a higher presence of points in regions B, A) and had medium contrast (higher or total presence of points in regions C) compared with the distributions descriptors using the proposed ACS.

The $\mu(\%)$ DL distributions of the original CMU-PIE data set before and after applying the proposed ACS along with the six contrast enhancement methods shown in Fig. 22.

The $\mu(\%)$ DL distributions from the original CMU-PIE images before enhancement, with distributions descriptors \underline{BACD} as shown in Fig. 22a. The $\mu(\%)$ in region B was

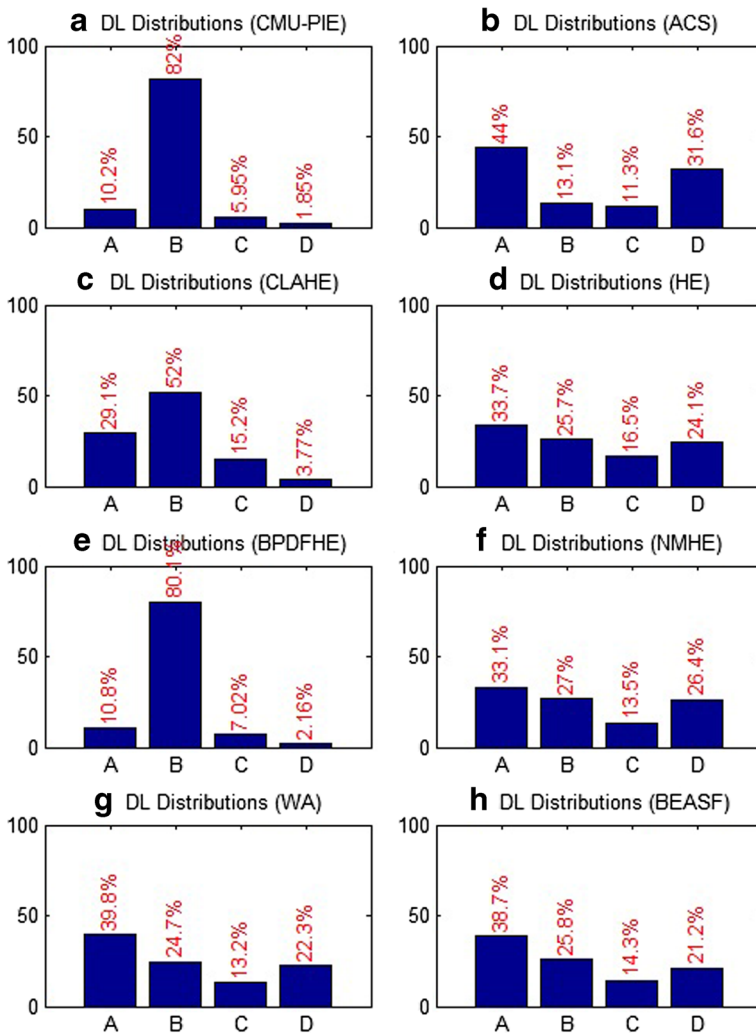


Fig. 22 $\mu(\%)$ of DL distributions at each region of the **a** original CMU-PIE data set before and after applying the seven contrast enhancement methods **b** ACS, **c** CLAHE, **d** HE, **e** BPDFHE, **f** NMHE, **g** WA, and **h** BEASF



Fig. 23 Visual interpretation of original face images for one subject in the EYale-B and the contrast enhanced with their directional gradient produced from *left to right* using: CLAHE, HE, BPDFHE, NMHE, WA, BEASF and ACS

higher compared to other regions, signified that the entire data set underwent dark and had poor quality before enhancement. Hence, the CMU-PIE data set considered as Low-Contrast/Low-Luminance (LC/LL).

Shown in Fig. 22b, a new significant distributions descriptors \underline{ADBC} occurred for CMU-PIE data set by employing ACS method. Thus, the points in regions D and A expanded, and the points in region B and C decreased, unlike the movements of the points when using the original distributions descriptor. Hence, the entire CMU-PIE data set images enhanced using the proposed ACS and generally considered as High-Contrast/Medium-Luminance (HC/ML).

Figure 22c, d, e, f, g, h, shown respectively the distributions descriptors of the six contrast enhancement methods (CLAHE, HE, BPDFHE, NMHE, WA and BEASF) with the

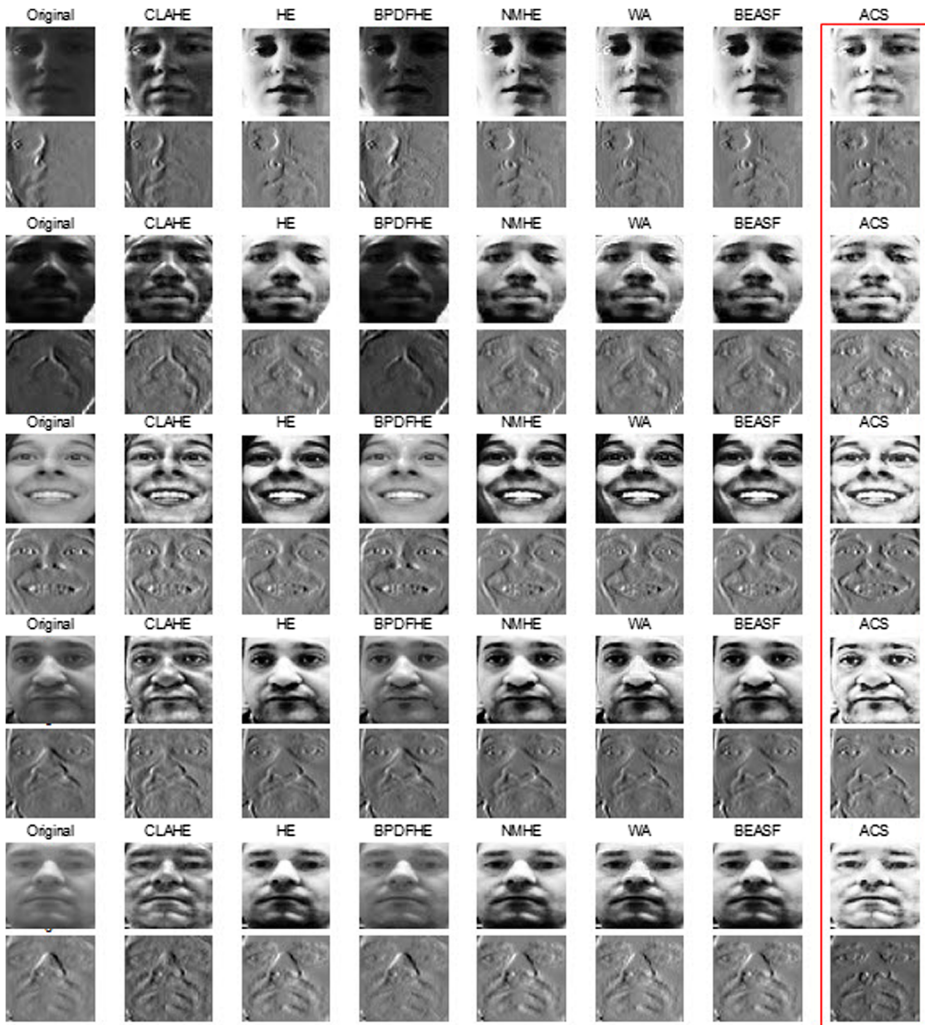


Fig. 24 Visual interpretation of original face images for different subject in the Mobio and the contrast enhanced with their directional gradient produced from *left to right* using: CLAHE, HE, BPDFHE, NMHE, WA, BEASF and ACS

DL distributions descriptors receptively, \underline{BACD} , \underline{ABDC} , \underline{BACD} , \underline{ABDC} , \underline{ABDC} , and \underline{ABDC} . In general, those $\mu(\%)$ had a new distributions descriptors occurred with major points concentrated in both regions B and A, signified that the entire CMU-PIE data set enhanced slightly with medium to lower luminance and had medium contrast compared with the distributions descriptors by employing the proposed ACS.

To summarise, the subjective order of all regions using the proposed ACS method is in identical order induced by the four data sets. These results demonstrate that the proposed method effectively enhance the face image while more structure details appears in all data sets. ACS method is also comparable to the six state-of-the-art contrast enhancement

methods with respect to a good contrast and uniform illumination as demonstrated by the identical order of the all regions induced.

3.3.3 Visual interpretation

Figures 23, 24, 25 and 26 shows the visual interpretation of face sample images for different subject respectively in the EYale-B, Mobio, Feret and CMU-PIE data sets and the contrast enhanced results with their directional gradient versions produced with the seven enhanced methods (CLAHE, HE, BPDFHE, NMHE, WA, BEASF and ACS). The visually results of face images clarify that ACS is motivated with a profitable enhancement rate in affording tremendous detail concerning face recognition systems.

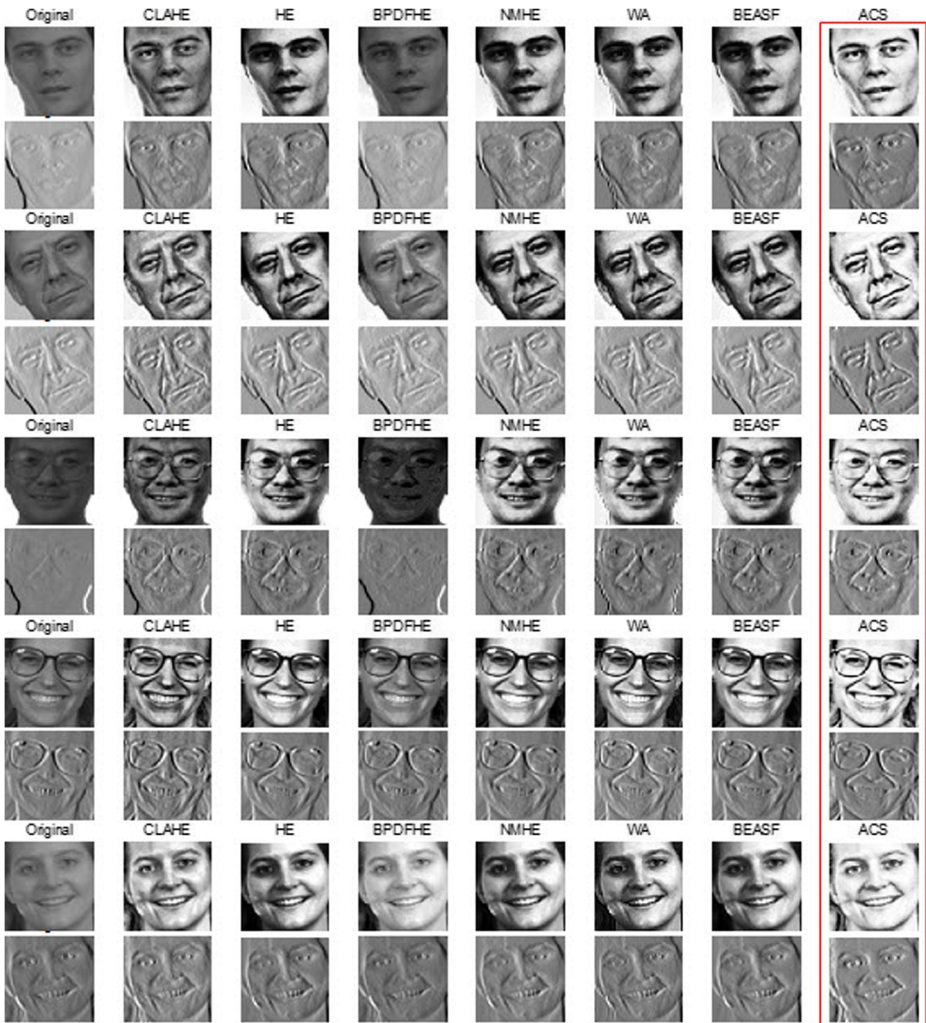


Fig. 25 Visual interpretation of original face images for different subject in the Feret and the contrast enhanced with their directional gradient produced from *left to right* using; CLAHE, HE, BPDFHE, NMHE, WA, BEASF and ACS



Fig. 26 Visual interpretation of original face images for one subject in the CMU-PIE and the contrast enhanced with their directional gradient produced from *left to right* using: CLAHE, HE, BPDFHE, NMHE, WA, BEASF and ACS

3.4 ACS performance through video scenario

The aim here is to evaluate the robustness of the proposed ACS within video surveillance scenario. Therefore, ChokePoint dataset [45] is selected in our experiment, which contains three cameras placed above a door, used for recording the entrance people from three viewpoints. Firstly, Viola and Jones algorithm [43] was applied for face detection in the video scene. Secondly, the detected face has been enhanced through the seven selected algorithms. Thirdly, The ACS performance was computed via IQM via BR-SM based on the average values from the selected video frames. Summarised detail of measurement results are given in Table 3, where the IQM via BR-SM of the ACS are proven to have

Table 3 Image quality measurement (IQM) via BR-SM of ChokePoint video surveillance

Video	Original	ACS	CLAHE	HE	BPDFHE	NMHE	WA	BEASF
ChokePoint								
EME	9.0129	28.713*	18.408	19.319	11.360	19.960	20.053	20.028
EMEE	40.239	45.542*	36.072	40.140	37.421	42.042	41.111	43.625
SDME	0.1087	0.2020*	0.1248	0.1709	0.1089	0.1678	0.1628	0.1678
SHP	0.0183	0.0676*	0.0481	0.0533	0.0256	0.0525	0.0530	0.0525
CPP	0.0357	0.0834*	0.0576	0.0677	0.0400	0.0679	0.0678	0.0679



Fig. 27 Visual interpretation of frames faces for one subject in the ChokePoint video surveillance and the enhancement produced from *left to right* using: CLAHE, HE, BPDFHE, NMHE, WA, BEASF and ACS

relatively the high mean value of EME, EMEE, SDME, SHP and CPP metrics compared to other methods which are found to have the almost the lowest measurements. In summary, these results based BR-SM indicates that the proposed ACS improve quite effectively the texture information of the detected facial regions. Illustrated in Fig. 27, the visual interpretation of frames faces resulting for one subject in the video scenario using the seven enhancement methods.

4 Conclusion

In this paper, a simple contrast stretching function operated adaptively towards nonuniform illumination face image enhancement was proposed. Contrast and illumination enhancement achieved by using frame differencing module (FDM) process to predict the pair (E, G), respectively based (i) E-generation via transformed enhancer (E-TE) and (ii) G-generation via DL descriptors (G-DL). The DL descriptor was examined to generate the face feature vectors. Subsequently, FDM was developed for faces label generations. Vocabulary tree model was examined to model DL descriptors and predict faces labels. The VTM performance was examined by referring to the ROC and AUC score. An objective evaluation was examined using four different face data sets and different image quality metrics showed that the proposed ACS can; (1) effectively enhances face contrast while affording uniform illumination result images to a large extent, (2) effectively improves the local face contrast for images with juxtaposed bright and dark regions and even very dark images, without the tendency to result in over-enhancement, and (3) comparable to several state-of-the-art methods with respect to contrast enhancement. Moreover, the evaluation study demonstrated that observers rated the results obtained via ACS as higher quality than those obtained via the other methods assessed in this study. In order to improve the ACS approach with higher performance, future work should be considered as followings; To employ fuzzy model as an alternative to the VTM model. To use face morphological operation process to improve the feature descriptors which will, in turn, improve the decision performance. In addition, to extend the proposed ACS within a video surveillance workbench.

Acknowledgements This research received funding from Ministry of Higher Education and Scientific Research (MHESR) and Centre de Développement des Technologies Avancées (CDTA)-Algeria, under the Science Fund Project (FNR-2013-2016).

References

1. Agaian S, Roopaei M, Akopian D (2014) Thermal-image quality measurements. In: 2014 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp 2779–2783. doi:[10.1109/ICASSP.2014.6854106](https://doi.org/10.1109/ICASSP.2014.6854106)
2. Agaian S, Roopaei M, Shadaram M, Bagalkot SS (2014) Bright and dark distance-based image decomposition and enhancement. In: 2014 IEEE international conference on imaging systems and techniques (IST) proceedings, pp 73–78. doi:[10.1109/IST.2014.6958449](https://doi.org/10.1109/IST.2014.6958449)
3. Agaian SS, Panetta K, Grigoryan AM (2000) A new measure of image enhancement. In: IASTED international conference on signal processing & communication, Citeseer, pp 19–22
4. Agaian SS, Panetta K, Grigoryan AM (2001) Transform-based image enhancement algorithms with performance measure. *IEEE Trans Image Process* 10(3):367–382
5. Arriaga-Garcia EF, Sanchez-Yanez RE, Garcia-Hernandez M (2014) Image enhancement using bi-histogram equalization with adaptive sigmoid functions. In: International conference on electronics, communications and computers (CONIELECOMP), 2014. IEEE, pp 28–34

6. Brown CD, Davis HT (2006) Receiver operating characteristics curves and related decision measures: a tutorial. *Chemom Intell Lab Syst* 80(1):24–38
7. Chang S-J, Li S, Andreassen A, Sha X-Z, Zhai X-Y (2015) A reference-free method for brightness compensation and contrast enhancement of micrographs of serial sections. *PloS one* 10(5):e0127855
8. Chen W, Er MJ, Wu S (2006) Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *IEEE Trans Syst Man Cybern B Cybern* 36(2):458–466
9. Choi Y, Kim H-I, Ro YM (2016) Two-step learning of deep convolutional neural network for discriminative face recognition under varying illumination. *Electronic Imaging* 2016(11):1–5
10. Eramian M, Mould D (2005) Histogram equalization using neighborhood metrics. In: *The 2nd Canadian conference on computer and robot vision (CRV'05)*. IEEE, pp 397–404
11. Faraji MR, Qi X (2014) Face recognition under varying illumination based on adaptive homomorphic eight local directional patterns. *IET Comput Vis* 9(3):390–399
12. Georghiades A, Belhumeur P, Kriegman D (2001) From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Trans Pattern Anal Mach Intell* 23(6):643–660
13. Gross R, Brajovic V (2003) An image preprocessing algorithm for illumination invariant face recognition. In: *International conference on audio-and video-based biometric person authentication*, Springer, pp 10–18
14. Hasikin K, Isa NAM (2012) Enhancement of the low contrast image using fuzzy set theory. In: *UKSim 14th international conference on computer modelling and simulation (UKSim)*, 2012, IEEE, pp 371–376
15. Heusch G, Cardinaux F, Marcel S (2005) Lighting normalization algorithms for face verification. Tech rep, IDIAP
16. Hu H (2015) Illumination invariant face recognition based on dual-tree complex wavelet transform. *IET Comput Vis* 9(2):163–173
17. Jobson DJ, Rahman Z-u, Woodell GA (1997) A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans Image Process* 6(7):965–976
18. Kryszczuk K, Drygajlo A (2006) On combining evidence for reliability estimation in face verification. In: *Signal processing conference, 2006 14th European*. IEEE, pp 1–5
19. La Cascia M, Sclaroff S, Athitsos V (2000) Fast, reliable head tracking under varying illumination: an approach based on registration of texture-mapped 3d models. *IEEE Trans Pattern Anal Mach Intell* 22(4):322–336
20. Lai Z-R, Dai D-Q, Ren C-X, Huang K-K (2015) Multiscale logarithm difference edgemaps for face recognition against varying lighting conditions. *IEEE Trans Image Process* 24(6):1735–1747
21. Lee K-C, Ho J, Kriegman DJ (2005) Acquiring linear subspaces for face recognition under variable lighting. *IEEE Trans Pattern Anal Mach Intell* 27(5):684–698
22. Liang H, Weller DS (2016) Comparison-based image quality assessment for selecting image restoration parameters. *IEEE Trans Image Process* 25(11):5118–5130
23. McCool C, Marcel S, Hadid A, Pietikainen M, Matejka P, Cernocky J, Poh N, Kittler J, Larcher A, Levy C, Matrouf D, Bonastre J-F, Tresadern P, Cootes T (2012) Bi-modal person recognition on a mobile phone: using mobile phone data. In: *IEEE ICME workshop on hot topics in mobile multimedia*
24. Mustapha A, Hussain A, Samad SA, Zulkifley MA (2014) Toward under-specified queries enhancement using retrieval and classification platforms. In: *IEEE symposium on computational intelligence for multimedia, signal and vision processing (CIMSIVP)*, 2014. IEEE, pp 1–7
25. Nikan S, Ahmadi M (2015) Local gradient-based illumination invariant face recognition using local phase quantisation and multi-resolution local binary pattern fusion. *IET Image Process* 9(1):12–21
26. Nister D, Stewenius H (2006) Scalable recognition with a vocabulary tree. In: *2006 IEEE Computer society conference on computer vision and pattern recognition (CVPR'06)*, vol 2. IEEE, pp 2161–2168
27. Oppenheim Av, Schafer R, Stockham T (1968) Nonlinear filtering of multiplied and convolved signals. *IEEE Trans Audio Electroacoust* 16(3):437–466
28. Park YK, Park SL, Kim JK (2008) Retinex method based on adaptive smoothing for illumination invariant face recognition. *Signal Process* 88(8):1929–1945
29. Phillips PJ, Wechsler H, Huang J, Rauss PJ (1998) The feret database and evaluation procedure for face-recognition algorithms. *Image Vision Comput*:295–306
30. Pizer S, Amburn E, Austin J, Cromartie R, Geselowitz A, Greer T, ter Haar Romeny B, Zimmerman J, Zuiderveld K (1987) Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing* 39(3):355–368
31. Poddar S, Tewary S, Sharma D, Karar V, Ghosh A, Pal SK (2013) Non-parametric modified histogram equalisation for contrast enhancement. *IET Image Process* 7(7):641–652
32. Restrepo A, Ramponi G (2008) Word descriptors of image quality based on local dispersion-versus-location distributions. In: *Signal processing conference, 2008 16th European*, IEEE, pp 1–5

33. Roopaiei M, Agaian S, Shadaram M, Hurtado F (2014) Cross-entropy histogram equalization. In: 2014 IEEE international conference on systems, man, and cybernetics (SMC), pp 158–163. doi:[10.1109/SMC.2014.6973900](https://doi.org/10.1109/SMC.2014.6973900)
34. Sheet D, Garud H, Suveer A, Mahadevappa M, Chatterjee J (2010) Brightness preserving dynamic fuzzy histogram equalization. *IEEE Trans Consum Electron* 56(4):2475–2480
35. Sheikh HR, Bovik AC (2006) Image information and visual quality. *IEEE Trans Image Process* 15(2):430–444
36. Sim T, Baker S, Bsat M (2002) The cmu pose, illumination, and expression (pie) database. In: Proceedings of the fifth IEEE international conference on automatic face and gesture recognition, FGR '02. IEEE Computer Society, Washington, DC, USA, p 53. <http://dl.acm.org/citation.cfm?id=874061.875452>
37. Sovierzoski MA, Argoud FIM, de Azevedo FM (2008) Evaluation of ann classifiers during supervised training with roc analysis and cross validation. In: 2008 international conference on biomedical engineering and informatics, vol 1. IEEE, pp 274–278
38. Tan X, Triggs B (2010) Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans Image Process* 19(6):1635–1650. doi:[10.1109/TIP.2010.2042645](https://doi.org/10.1109/TIP.2010.2042645)
39. Thamizharasi A, Jayasudha J (2016) An illumination invariant face recognition by enhanced contrast limited adaptive histogram equalization. *ICTACT Journal on Image & Video Processing* 6(4)
40. Tilbury JB, Van Eetvelt W, Garibaldi JM, Curnsw J, Ifeachor EC (2000) Receiver operating characteristic analysis for intelligent medical systems—a new approach for finding confidence intervals. *IEEE Trans Biomed Eng* 47(7):952–963
41. Tizhoosh HR (2000) Fuzzy image enhancement: an overview. In: *Fuzzy techniques in image processing*, Springer, pp 137–171
42. Venkateshwarlu K (2010) Image enhancement using fuzzy inference system. Ph.D. thesis, Thapar University Patiala
43. Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition, 2001. CVPR 2001, vol 1. IEEE, pp I–I
44. Štruc V, Pavešić N (2009) Illumination invariant face recognition by non-local smoothing. In: *European workshop on biometrics and identity management*, Springer, pp 1–8
45. Wong Y, Chen S, Mau S, Sanderson C, Lovell BC (2011) Patch-based probabilistic image quality assessment for face selection and improved video-based face recognition. In: *IEEE computer society conference on computer vision and pattern recognition workshops (CVPRW)*, 2011. IEEE, pp 74–81
46. Zhou Y, Panetta K, Agaian S (2010) Human visual system based mammogram enhancement and analysis. In: *2nd international conference on image processing theory tools and applications (IPTA)*, 2010. IEEE, pp 229–234
47. Zuiderveld K (1994) Contrast limited adaptive histogram equalization



Aouache Mustapha received the B.Sc. in control and automation from Boumerdes University, Algeria; M.Sc and PhD in electrical, electronic and system engineering from Universiti Kebangsaan Malaysia (UKM), respectively. Currently, he is a senior research fellow at the Telecom Division, Centre de Développement des Technologies Avancées (CDTA), Algeria. His current research interests include the control system, intelligent systems, image processing and diagnostic systems, CBIR and CADs system for the medical application.



Adel Oulefki received his M.Sc and PhD degrees in Electrical Engineering in 2009 and 2015 respectively, University of Bordj Bou Arréridj, Algeria. He worked as a Postdoctoral Fellow at MSE Laboratory from 2014 to 2015. He has been a temporary lecturer in the Department of Electronics, University of Bordj Bou Arréridj from 2013 to 2014. Currently, he is a senior research associate at the Centre de Développement des Technologies Avancées (CDTA), Algeria. His current research involves data clustering, signal, image and video processing, real-time pattern recognition, reconfigurable architectures and biometric authentication.



Messaoud Bengherabi received the state engineer degree in electrical engineering option communication from the National Institute of Electricity and Electronics (INELEC), Boumerdes, Algeria in 1996 and the Magister degree in advanced signal processing from the Ecole Polytechnique (EMP), Bordj El Bahri, Algeria, in 1999. Since joining the Centre de Développement des Technologies Avancées (CDTA) in October 1999, he has been engaged in research on low delay narrowband-wideband CELP-based speech coders and biometrics. Currently, he is the head of the Biometric team in CDTA. His research interests include speaker and face recognition, multimodal biometrics and forensic applications of biometric systems.



Elhocine Boutellaa received both Magister and PhD degrees in computer science from Ecole nationale Supérieure d'Informatique (ESI), Algeria, in 2011 and 2017 respectively. He has been working as research associate with Centre de Développement des Technologies Avancées (CDTA) from 2010 to 2016. He is currently a senior researcher fellow with CDTA. His research interests include face analysis, biometrics, computer vision and deep learning.



Mustafa Almahdi Algaet was born in Al Khums, Libya. In 2004 earned bachelor's degree in electrical engineering from Almahdi University of Libya and earned master's degree in information and Communication technology from University Utara Malaysia (UUM), in (2006). Currently, he is a Ph.D candidate in computer science from Universiti Teknikal Malaysia Melaka (UTeM) in Malaka, Malaysia. His research interests include computer vision and wireless telemedicine for e-health services.