

An Illumination Invariant Accurate Face Recognition with Down Scaling of DCT Coefficients

Virendra P. Vishwakarma, Sujata Pandey and M. N. Gupta

Department of Computer Science and Engineering, Amity School of Engineering and Technology, New Delhi, India

In this paper, a novel approach for illumination normalization under varying lighting conditions is presented. Our approach utilizes the fact that discrete cosine transform (DCT) low-frequency coefficients correspond to illumination variations in a digital image. Under varying illuminations, the images captured may have low contrast; initially we apply histogram equalization on these for contrast stretching. Then the low-frequency DCT coefficients are scaled down to compensate the illumination variations. The value of scaling down factor and the number of low-frequency DCT coefficients, which are to be rescaled, are obtained experimentally. The classification is done using k -nearest neighbor classification and nearest mean classification on the images obtained by inverse DCT on the processed coefficients. The correlation coefficient and Euclidean distance obtained using principal component analysis are used as distance metrics in classification. We have tested our face recognition method using Yale Face Database B. The results show that our method performs without any error (100% face recognition performance), even on the most extreme illumination variations. There are different schemes in the literature for illumination normalization under varying lighting conditions, but no one is claimed to give 100% recognition rate under all illumination variations for this database. The proposed technique is computationally efficient and can easily be implemented for real time face recognition system.

Keywords: discrete cosine transform, correlation coefficient, face recognition, illumination normalization, nearest neighbor classification

1. Introduction

Two-dimensional pattern classification plays a crucial role in real-world applications. To build high-performance surveillance or information security systems, face recognition has been known as the key application attracting enormous researchers highlighting on related topics

[1,2]. Even though current machine recognition systems have reached a certain level of maturity, their success is limited by the real applications constraints, like pose, illumination and expression. The FERET evaluation shows that the performance of a face recognition system decline seriously with the change of pose and illumination conditions [31].

To solve the variable illumination problem a variety of approaches have been proposed [3, 7-11, 26-29]. Early work in illumination invariant face recognition focused on image representations that are mostly insensitive to changes in illumination. There were approaches in which the image representations and distance measures were evaluated on a tightly controlled face database that varied the face pose, illumination, and expression. The image representations include edge maps, 2D Gabor-like filters, first and second derivatives of the gray-level image, and the logarithmic transformations of the intensity image along with these representations [4].

The different approaches to solve the problem of illumination invariant face recognition can be broadly classified into two main categories. The first category is named as *passive* approach in which the visual spectrum images are analyzed to overcome this problem. The approaches belonging to other category named *active*, attempt to overcome this problem by employing active imaging techniques to obtain face images captured in consistent illumination condition, or images of illumination invariant modalities. There is a hierarchical categorization of these two approaches. An extensive review of both approaches is given in [5].

Different methods of *passive* approach can be broadly classified into three main categories. The first approach is based on “preprocessing and normalization”. The representative methods are histogram equalization, Gamma correction, logarithm transform, etc. for illumination normalization [4]. But these global processing techniques of image processing are found to be insufficient to overcome variations due to illumination changes. Bischof et al. [32] devised a method to integrate a set of gradient based filter banks into the eigenspace framework, for the purpose of illumination insensitive object recognition. This method concatenates the benefits of global eigenspace representations and local representations through filter banks. The eigen-image coefficients were calculated by solving a system of linear equations in place of projection. These eigen-image coefficients are invariant for linearly filtered images. This results in recognition improvement on an object database comprising varying pose of objects under a point light source [32]. But the method is silent about illumination normalization and compensation.

In the second approach, the systems are exploiting “invariant feature extraction” method. A well established method for feature extraction is Fisher-face (based on linear discriminant analysis (LDA)), which maps the image space to a low dimensional subspace to ignore the variations in lighting etc. [7]. This method is a statistical linear projection method in which the representativeness of the training samples controls the performance of the system. An illumination invariant signature image is generated using a bootstrap set [8]. This signature image is viewed as a quotient image, which can be used for face recognition under varying lighting conditions. This method needs a bootstrap database and the performance degrades when dominant features between the bootstrap set and the test set are misaligned.

The third approach deals with face modeling methods which explores 3-D shape of human faces. Here the attempt is to construct a generative 3-D face model that can be used to render the face images with different poses and illuminations [3,9-11]. A generative model called illumination cone was presented in [3,9]. This method shows that an illumination convex cone can be constructed from a set of face images

captured in fixed pose, but under different illumination conditions. A low-dimensional linear subspace can be obtained by approximating this illumination cone. The measure drawback of model based approaches is that a training set containing face images of a subject under different lighting conditions is needed. This limits the application of this type of approach in practical face recognition system. Our approach of illumination normalization belongs to “preprocessing and normalization” technique for compensating illumination variations. In this paper, we propose a new approach that utilizes the image enhancement capability of histogram equalization and DCT. The input image contrast stretching is done by histogram equalization. To suppress illumination variations DCT is applied on the resultant image. As the low-frequency DCT coefficients correspond to illumination variations, the authors propose a new technique to suppress them. The method involves the down scaling of these coefficients. The value of rescaling down factor and the number of low-frequency DCT coefficients, which are to be rescaled, are obtained experimentally. Further, the mean and k-way nearest neighbor classification is applied on the processed image. The results show accurate recognition in all illumination variation conditions of the database used.

The reminder of this paper is organized as follows. The brief description of histogram equalization, DCT, principle component analysis, correlation and nearest neighbor classification are given in Section II. System Architecture and details of database used, are presented in Section III. In Section IV, experimental results and discussions are carried out and the conclusions are presented in Section V.

2. Preliminaries

2.1. Histogram Equalization

Histogram equalization maps the input image’s intensity values so that the histogram of the resulting image will have an approximately uniform distribution [12-14]. The histogram of a digital image with gray levels in the range $[0, L-1]$ is a discrete function

$$p(r_k) = \frac{n_k}{n} \quad (1)$$

where L is the total number of gray levels (for 8-bit gray level representation, L is equal to 256), r_k is the k^{th} gray level, n_k is the number of pixels in the image with that gray level, n is the total number of pixels in the image, and $k = 0, 1, 2, \dots, L - 1$. $p(r_k)$ gives an estimate of the probability of occurrence of gray level r_k .

By histogram equalization, the local contrast of the object in the image is increased, especially when the usable data of the image is represented by close contrast values. Through this adjustment, the intensity can be better distributed on the histogram. This allows for areas of lower local contrast to gain a higher contrast without affecting the global contrast. The histogram equalization for a digital image is defined as a transformation on the input intensity levels (r_k) to obtain output intensity levels (s_k) as

$$s_k = T(r_k) = \sum_{j=1}^k p_r(r_j) = \sum_{j=1}^k \frac{n_j}{n} \quad (2)$$

for $k = 0, 1, 2, \dots, L - 1$.

2.2. Discrete Cosine Transform

The DCT is a popular technique in imaging and video compression, which transforms signals from the spatial representation into a frequency representation. The forward 2D-DCT [13-15] of a $M \times N$ block image is defined as

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \times \cos\left[\frac{\pi(2x+1)u}{2M}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right]. \quad (3)$$

The inverse transform is defined as

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u, v) \times \cos\left[\frac{\pi(2x+1)u}{2M}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (4)$$

where

$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{M}}; & u = 0 \\ \sqrt{\frac{2}{M}}; & u = 1, 2, \dots, M - 1 \end{cases},$$

$$\alpha(v) = \begin{cases} \frac{1}{\sqrt{N}}; & v = 0 \\ \sqrt{\frac{2}{N}}; & v = 1, 2, \dots, N - 1 \end{cases},$$

x and y are spatial coordinates in the image block, u and v are coordinates in the DCT coefficients block. Figure 1 shows the properties of the DCT coefficients in $M \times N$ blocks with the zigzag pattern used by JPEG compression to process the DCT coefficients. Although the total energy remains the same in the $M \times N$ blocks, the energy distribution changes with most energy being compacted to the low-frequency coefficients. In our approach, we have taken the DCT coefficients in zigzag pattern as shown in Figure 1.

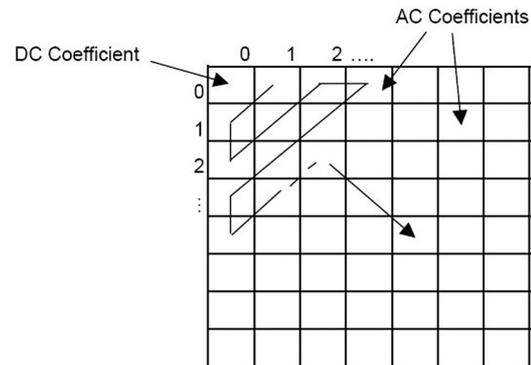


Figure 1. Block feature of DCT coefficients and their selection in zigzag pattern.

The DC coefficient is represented by $C(0, 0)$ in the forward 2D-DCT equation. As the cosine of zero is one, the equation is simplified to:

$$C(0, 0) = \frac{1}{\sqrt{MxN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y). \quad (5)$$

The DC coefficient, which is located at the upper left corner, holds most of the image energy and represents the proportional average of the

$M \times N$ blocks. The remaining $((M \times N) - 1)$ coefficients denote the intensity changes among the block images and are referred to as AC coefficients. The DCT is performed on the entire image obtained after processing the input face images by histogram equalization in our approach.

2.3. Correlation Coefficient

The correlation coefficient [13, 16-18] between two images $A(x, y)$ and $B(x, y)$ is defined as

$$\text{CorrCoef} = \frac{\sum_s \sum_t [A(s, t) - \bar{A}][B(s, t) - \bar{B}]}{\left\{ \sum_s \sum_t [A(s, t) - \bar{A}]^2 \sum_s \sum_t [B(s, t) - \bar{B}]^2 \right\}^{\frac{1}{2}}}. \quad (6)$$

The summation is taken over the whole image region, as we are taking the images A and B of same size. \bar{A} and \bar{B} are the average values of images A and B respectively. As correlation techniques are computationally expensive and require great amounts of storage, it is natural to pursue dimensionality reduction schemes. In place of using complete face image for calculating correlation coefficient, a simple variant of this is used in [16] for face space dimension reduction. Here in place of full-face image, the eyes, nose and mouth regions of a face image are used. This technique gives about only 50% of dimension reduction. We have used both techniques for finding the correlation coefficient between two face-images in our experiments, which is further described in the next section.

2.4. Principal Component Analysis

One of the most popular schemes of dimension reduction, is Principal Component Analysis (PCA) [7,9,19,20]. We have implemented PCA procedure in our experiments as described by Turk and Pentland [19]. In a training set of M face images, let a face image be represented as a two dimensional N by N array of intensity values, or a vector of dimension N^2 . Then PCA tends to find a M' -dimensional subspace whose basis vectors correspond to the maximum variance direction in the original image space. This

new subspace is normally lower dimensional ($M' < M \ll N^2$). New basis vectors define a subspace of face images called face space. All images of known faces are projected onto the face space to find sets of weights that describe the contribution of each vector. To identify an unknown image, the image is projected onto the face space as well to obtain its set of weights. By comparing a set of weights for the unknown face to sets of weights of known faces, the face can be identified. If the image elements are considered as random variables, the PCA basis vectors are defined as eigenvectors of the scatter matrix S defined as:

$$S = \sum_{i=1}^M (x_i - \mu) \cdot (x_i - \mu)' \quad (7)$$

where μ is the mean of all images in the training set (we have taken Subset1 as the training set) and x_i is the i^{th} face image represented as a vector i. e. its columns are concatenated in a vector form. A projection matrix W is generated using M' eigenvectors corresponding to M' largest eigenvalues of S , thus creating a M' -dimensional face space. As this face space is generated using eigenvectors of scatter matrix, sometimes this is also called as eigenspace.

2.5. Nearest Neighbor Classification

Classification (similarity search) is a very crucial step in any pattern recognition application. One of the most popular non-parametric techniques is the Nearest Neighbor classification (NNC), which simply states that the unclassified object is assigned to the class of its Nearest Neighbor (NN) among a set of design objects. NNC asymptotic or infinite sample size error is less than twice of the Bayes error [21]. Various variants of NNC are given in the literature [22-25]. In our implementation, we have used k -Nearest Neighbor Classifier (k -NNC) and Nearest Mean Classifier (NMC). The basic NNC rule behind these techniques is given by Cover and Hart [21]. In k -NNC, instead of 1-NN, generally, k -nearest neighboring data objects are considered. Then, the class label of unseen objects is established by majority vote. The parameter k represents the number of neighbors involved. Tuning k as a way to regularize the NNC gives

a trade-off between the distribution of the training data with a priori probability of the classes involved. When $k = 1$, the training data distribution and a priori probability are considered, while, when $k = N$, only a priori probability of the classes determines the class label. In NMC, we consider only the mean of each class, i. e. one prototype per class. In comparison to NNC, NMC has a high error on the training data and on the test data, but the error on the training data is a good prediction of the error on the test data. When considered as a regularized version of the NNC with one NN, i. e. 1-NNC, NMC has only one prototype per class instead of as many prototypes as the number of training objects. Thus it is evident that reducing the number of labeled prototypes is another way of regularizing the 1-NNC, where a high number of prototypes makes the classifier more (training data) specific and a low number makes it more general [23]. In our implementation, as we have used both k -NNC and NMC techniques for classification, thus the performance of our approach for illumination normalization under varying lighting conditions, is tested in both extremes of NNC.

3. System Architecture and Database Used

3.1. System Architecture

The proposed implementation architecture is drawn in Figure 2. Here we are applying our new technique for illumination normalization under varying lighting conditions on both training database and test database. These databases are taken from Yale Face Database B. The description of this database is given in the next subsection. In illumination normalization, the histogram equalization (HE) is applied on the input images for contrast stretching. When the light source is not in front of the subject, the image captured in this condition has illumination variation. As the image has large illumination variation, i.e. the light source is more inclined, the image has lower contrast. This is reduced by HE. On the output of HE, DCT is applied to switch into the frequency domain from space domain. Here (in frequency domain) illumination variation is compensated by scaling down the low frequency DCT coefficients. The inverse DCT is applied on the processed

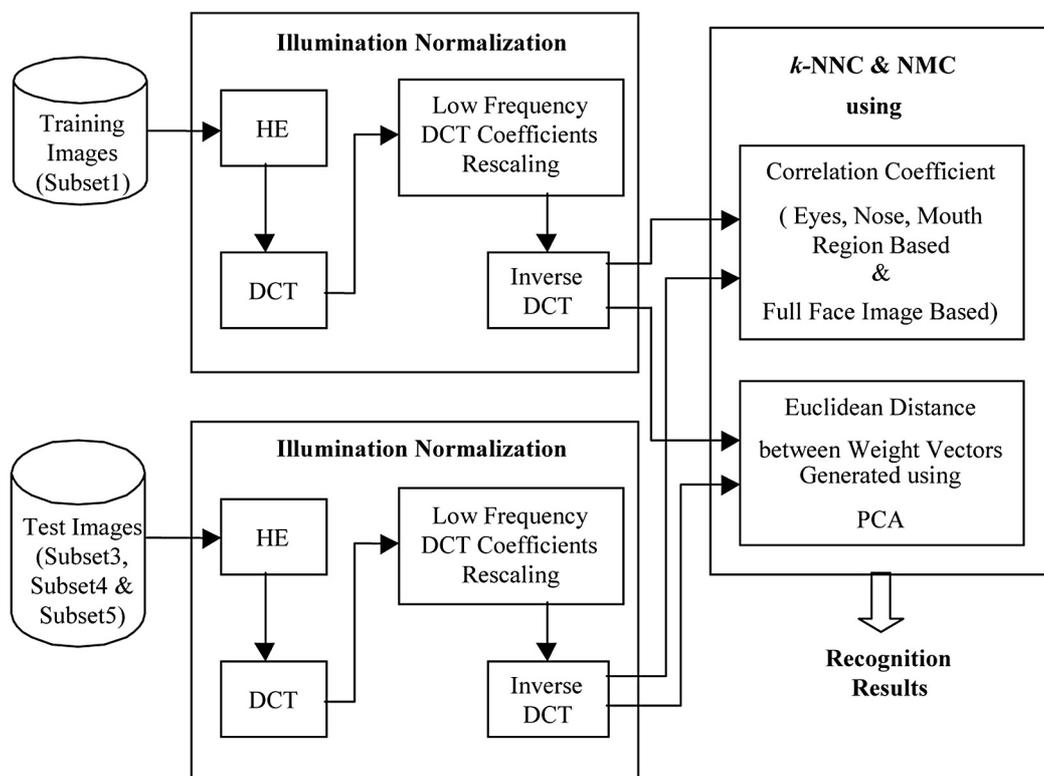


Figure 2. The system architecture of our approach for face recognition.

Subset	1	2	3	4	5
Azimuth & Elevation ($^{\circ}$)	0 ~ 12	13 ~ 25	26 ~ 50	51 ~ 77	> 77
Number of subset images	60	120	120	140	190

Table 1. Subsets of Yale Face Database B.

DCT coefficients to generate the illumination-normalized images. These images are fed to the classifier engine.

We have adopted two variants of classifiers, k -NNC and NMC. The distance metrics used in the classifiers are correlation coefficient and Euclidean distance. In k -NNC, all images of the Subset1 of Yale Face Database B are considered as training images, whereas, in NMC, the mean images corresponding to each subject generated using this subset (Subset1) are considered as training images. The calculation of correlation coefficient is done in two different manners. In one way, we have considered only eyes, nose and mouth region of the face image and in the other way, the full-face image is considered. The correlation coefficient is taken as the similarity score between two images and maximum value of this among different subjects is taken to find the nearest neighbor of a test image. The Euclidean distance is taken as the distance metric for the analysis in the case of PCA. First the eigenspace W is generated for the training images using M' eigenvectors of scatter matrix S . After generating eigenspace W , all images, the training as well as the test images are projected on W to generate the set of weights. Euclidean distances are calculated between weights corresponding to the unknown face (test set) and known faces (training set).

The minimum of these Euclidean distances is checked to find the nearest neighbor of a test image. After checking the correctness of nearest neighbors/mean in all cases, the recognition results are generated. Further elaborations of different steps of implementation are carried out in proceeding sub-sections.

3.2. Face Image Database Used

The experiments are carried out on the Yale Face Database B [30] which contains 5760 single light source images of 10 subjects, each seen under 576 viewing conditions (9 poses x 64 illumination conditions). As we are analyzing the

illumination normalization in this paper, only fixed frontal face images under varying lighting conditions are considered. These images are categorized in five subsets based on the azimuth and elevation of the light source direction with the camera's axis [9]. The azimuth and elevation of light source direction for these subsets are given in Table 1. All images were cropped to include only the face with as little hair and background as possible. Each frontal pose image was aligned (scaled and rotated) so that the eyes lay at a fixed distance apart (equal to four sevenths of the cropped window width) and on an imaginary horizontal line. Also, the face was centered along the vertical direction so that the two imaginary horizontal lines passing through the eyes and mouth were equidistant from the center of the cropped window. This alignment was performed in order to remove any bias from the recognition results due to the association of a particular scale, position, or orientation to a particular face [9]. The original size of the images is 640×480 pixels. After cropping these are of the resolution of 192×168 pixels. In our experiments, these images are further sub-sampled by 1.6 to have a resolution of 120×105 pixels. We have used Subset1 as the training set and all other four subsets as the test sets.

4. Experimental Results and Discussions

4.1. Histogram Equalization

As aforementioned, the image captured under varying illumination may have lower contrast. For contrast increasing, we apply histogram equalization (HE) on input images. The HE stretches the contrast of the input images to full range of intensity levels. Figure 3(a) shows one of the images taken from database and its corresponding histogram is shown in Figure 3(c). Figure 3(b) shows the processed image after applying histogram equalization on the image of Figure 3(a). The histogram of processed image is shown in Figure 3(d). The histogram of original image, as shown in Figure 3(b), depicts

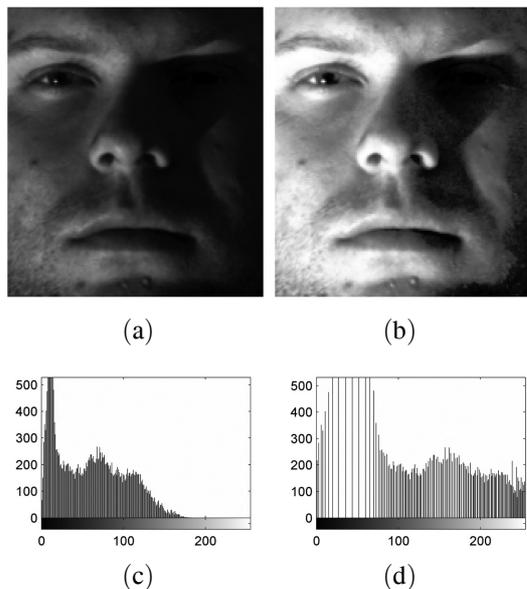


Figure 3. (a) Original Image; (b) Image after histogram equalization; (c) Histogram of image in (a); (d) Histogram of image in (b).

that the probability of occurrence of intensity levels is below the mid of gray scale. After histogram equalization, we obtain approximately uniform intensity distribution and we get a high contrast image as shown in Figure 3(b). But it is also visible from Figure 3(a) and Figure 3(b) that the illumination variations are not affected, these remain in the images, although illumination variations are shifted to the upward in the gray scale. This problem is solved by our new approach described in the next sub-section.

4.2. Illumination Normalization Using Rescaling of Low-frequency DCT Coefficients

Our approach for illumination normalization utilizes the fact that low-frequency DCT coef-

ficients are correlated with illumination variations. As mentioned in Figure 1 the DCT coefficients are selected in zig-zag pattern, if we select DCT coefficients from upper diagonal elements of top-left corner of Figure 1, the DCT coefficients will be in increasing order of frequency. For the resolution of images taken in the analysis, the total number of DCT coefficients are $120 \times 105 = 12600$. Out of these, the coefficients corresponding to low frequencies are selected for re-scaling. Now the questions arise what is the value of Rescaling Down Factor (*RDF*) and how many low-frequency DCT coefficients (C_{resc}) are to be rescaled. In next sub-sections, we are finding the dependence of performance variation with respect to the various values of *RDF* and C_{resc} . Figure 4 shows the change in appearance of the face-image shown in Figure 3(a) for increasing values of C_{resc} . The Figure 4 (a) shows the result of our method of illumination normalization with $C_{resc} = 5$. Similarly the subsequent images of Figure 4 are the result for different values of C_{resc} . For larger values of C_{resc} , the appearance of images shown in Figure 4 is not good. But there are sufficient numbers of high frequency facial features to do discrimination among the subjects. It will be found in sub-section '4.4' and '4.6.' that the best performance (accurate face recognition) is found for $C_{resc} = 37$, which corresponds to the processed image in Figure 4(e).

These values of C_{resc} are not the actual selected number of low-frequency DCT coefficients; this is the number of elements in the diagonal of a square matrix. In our implementation, we use this value (C_{resc}) to select the upper left diagonal elements of a matrix, thus this gives a reduced measure of low-frequency DCT coefficients selected. The actual number of low-frequency DCT coefficients, selected for rescal-

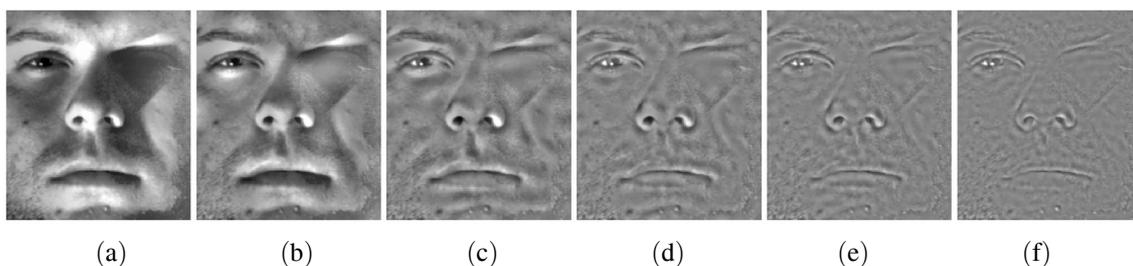


Figure 4. Illumination normalized images with different C_{resc} : (a) $C_{resc} = 5$; (b) $C_{resc} = 13$; (c) $C_{resc} = 21$; (d) $C_{resc} = 29$; (e) $C_{resc} = 37$; (f) $C_{resc} = 45$.

ing, can be calculated using following equation:

$$\text{Total number of low-frequency DCT coefficients} \\ = C_{resc} \times (C_{resc} - 1) - 1 \quad (8)$$

Eq. (8) gives the value of selected number of AC coefficients only. The first coefficient (DC coefficient $C(0,0)$) is not selected for re-scaling down, but it is enhanced. As for illumination normalization, we divide some (depends on the value C_{resc}) low-frequency DCT coefficients by RDF , when we apply the inverse DCT, the contrast of the image obtained is lower. Then the appearance of the processed image is darker. To reduce this effect, before applying inverse DCT, we increase the $C(0,0)$ by some factor (say 20%). This improves the visual appearance of the processed image. Although in experimentation, we found that the scaling up or down of $C(0,0)$ does not effect the recognition results. This is the reason that correlation coefficient calculation, inherently does image normalization. Similarly, in PCA also, before finding the face space, we are doing image normalization.

4.3. Experimentation Using Correlation Coefficient with Rescaling Down Factor

We have calculated the correlation coefficient in two different ways, in first method: eyes, nose and mouth region based and in second method: full-face image based. The Figure 5 shows the image regions of a face image selected for the calculation of correlation coefficient for first and second method.

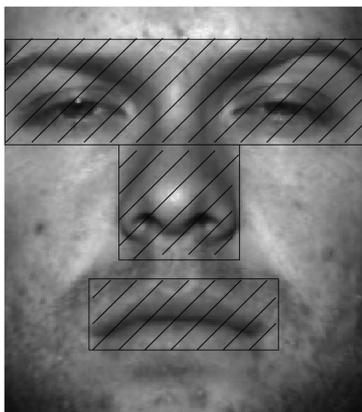


Figure 5. Eyes, nose and mouth region of a face image considered for correlation coefficient calculation.

In the first method, the hatched portion of face image are taken, which corresponds to the eyes, nose and mouth corresponding to the main dominant features of a subject in the face image. While in the second method, the complete face image shown in Figure 5 is considered. Based on these, the correlation coefficients are calculated to obtain the performance results.

Haped and Levine [6] have utilized feature extraction capabilities of DCT. They used certain subset of the DCT coefficients as a feature vector describing a face image. This feature vector contains the low-to-mid frequency DCT coefficients, as these are the ones having the highest variance. They applied this method of feature extraction on database of small illumination variations. For illumination compensation, they applied histogram equalization technique only [6]. As aforementioned (Figure 4), the illumination variations mainly correspond to low-frequency DCT coefficients. Thus illumination normalization means to obtain the low-frequency DCT coefficients which correspond to the uniform lighting as well as to preserve the low-frequency details of a subject. In our approach of illumination normalization, we have taken the concept of re-scaling of low-frequency DCT coefficients to a lower value. The idea of our approach of illumination normalization is inspired by the relation between low-frequency DCT coefficients of a highly illuminated face image and by that of a uniformly illuminated face image. For example, the value of (0,1)th DCT coefficient of a half-lighted face image is in the order of thousand, whereas that of a uniformly illuminated face image is in tens. Thus dividing the low-frequency DCT coefficients by a constant (named rescaling down factor (RDF) in our implementation) does this objective that eliminates the effect of illumination variation while it preserves the facial features corresponding to the low-frequencies.

We have done extensive experimentation to find the effect of RDF value variation on recognition error. Figure 6(a) shows the performance variation with respect to RDF for k -NNC and that for NMC is shown in Figure 6(b). We found that the recognition error is zero when RDF is equal to 15 to 35 for k -NNC. This also happens for NMC when RDF is equal to 9 to 25. Thus

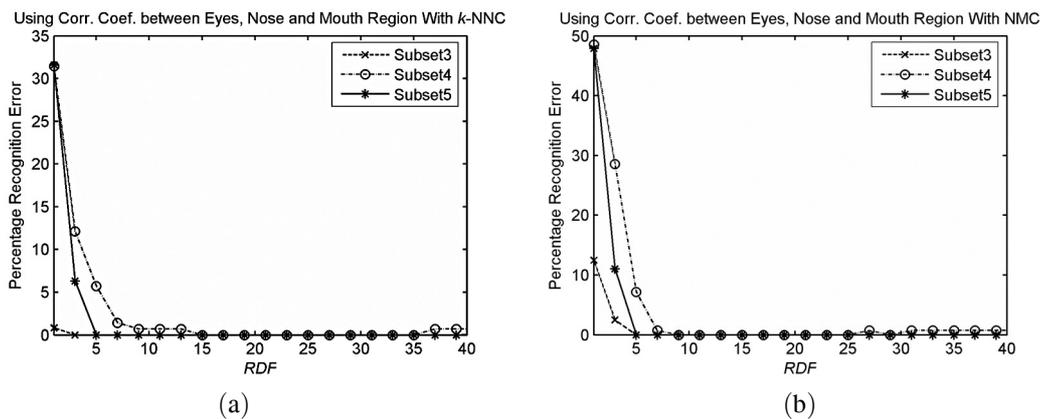


Figure 6. Performance variation using corr. coef. between eyes, nose and mouth region with different RDF for (a) k -NNC (b) NMC.

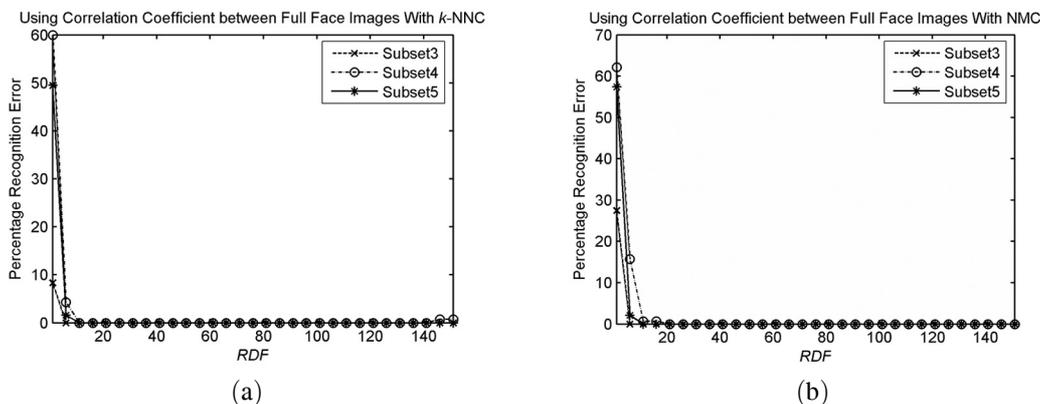


Figure 7. Performance variation using corr. coef. between full face image with different RDF for (a) k -NNC (b) NMC.

the mean value of RDF is 25 in case of k -NNC and that is 17 for NMC. Again, taking the mean value of 25 and 17, which is 21 as the value of RDF for further analysis to obtain the performance variation with respect to other parameter (C_{resc}) discussed in the next sub-section. The value of C_{resc} is 37 for the performance analysis shown in Figure 6. The graphs of Figure 6 also proves that our concept of re-scaling of low-frequency DCT coefficients is correct, as we found that for above mentioned range of RDF , the recognition error is zero and for the value of RDF above and below this range, the recognition rate is not zero.

Similarly, Figure 7 shows the performance variation using the second method of correlation co-

efficients calculation. In this method, the size of pattern vector corresponds to the full-face image based and this is larger in comparison to the size of pattern vector corresponding to eyes, nose and mouth region based. This is evident from Figure 7 that the range of RDF for which the recognition error is zero, is from 11 to 141 for k -NNC. In NMC, if RDF is above 21, the recognition error is zero. Considering k -NNC, we have mean value of RDF as 76. This value of RDF is considered for analyzing the performance variation with respect to C_{resc} when the correlation coefficients are calculated based on full-face image. In the analysis of Figure 7 also, the value of C_{resc} is taken as 37, as in the analysis shown in Figure 6.

4.4. Experimentation Using Correlation Coefficient with Number of Low Frequency DCT Coefficients

Here, the analysis of performance variation with respect to the number of low-frequency DCT coefficients (C_{resc}) is carried out. As in analysis using *RDF*, we are here also calculating the correlation coefficients in two ways, with eyes, nose and mouth region based and with full-face image based. The value of *RDF* is 21 in the first case and 76 in the second case, as mentioned in the previous sub-section. As evident from Figure 8 and Figure 9, the recognition error decreases significantly after a small number of coefficients selected for re-scaling. The recognition error is zero for C_{resc} ranging from 37 to 39 in case of eyes, nose and mouth region based correlation coefficients calculation (Figure 8). Similarly, the recognition error is

zero when C_{resc} is equal to 37 in correlation coefficients calculation based on full-face image. In both cases of classification, i. e. *k*-NNC and NMC the values of C_{resc} remain same. The difference lies in the performance error for the initial values of C_{resc} . When no illumination normalization is applied, i. e. when $C_{resc} = 1$, the recognition errors for different subsets using *k*-NNC are less than the recognition errors using NMC. As stated in system architecture sub-section, the subset1 of ‘Yale Face Database B’ is used as training set. The parameter *k* in *k*-nearest neighbor classifier is equal to six, since there are six images per subject in this subset, while in the nearest mean classifier a mean image is generated using these six images of a subject. There are six images with slight illumination variation to match with a query image in *k*-NNC and there is only one image (mean image) to match in NMC. Thus the probability

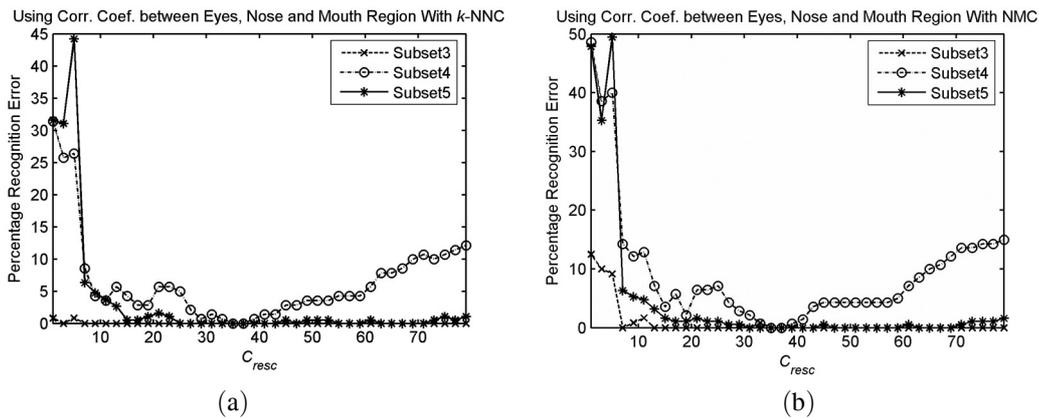


Figure 8. Performance variation using corr. coef. between eyes, nose and mouth region with different C_{resc} for (a) *k*-NNC (b) NMC.

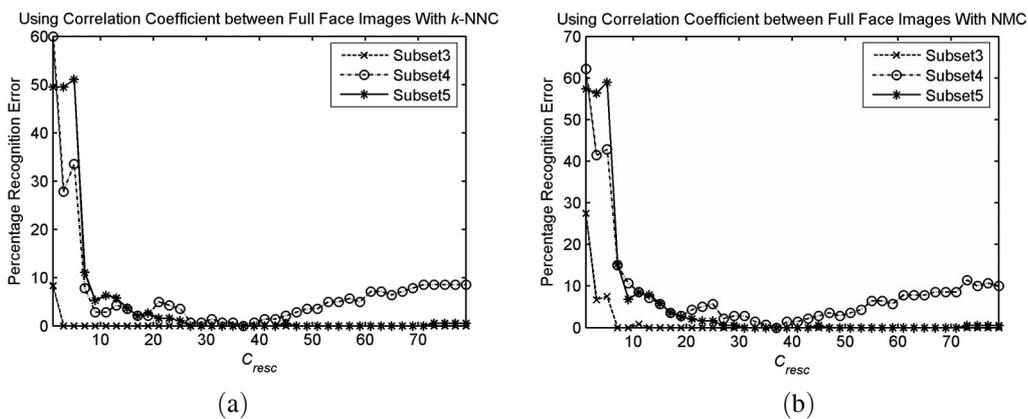


Figure 9. Performance variation using corr. coef. between full face image with different C_{resc} (a) *k*-NNC (b) NMC.

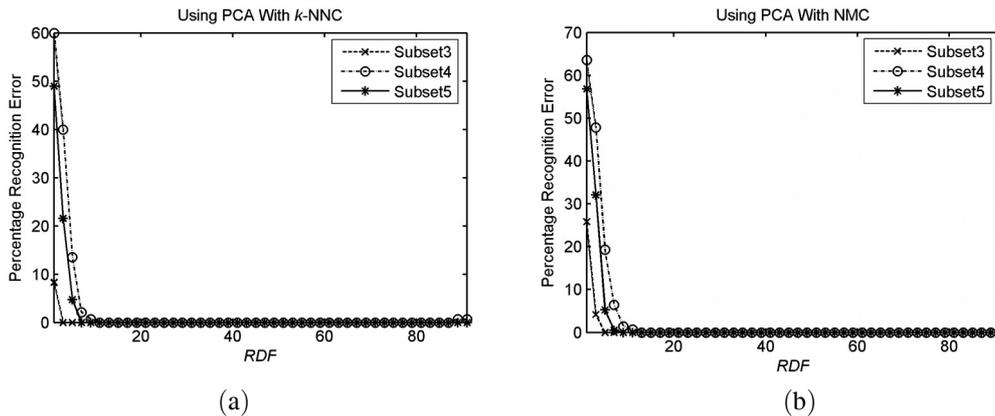


Figure 10. Performance variation using PCA with different RDF for (a) k -NNC (b) NMC.

of matching in k -NNC is greater in comparison to that in NMC. That is why we are getting less recognition error in k -NNC. All our implementation results, including previous and next sub-section results, are listed in Table 2.

4.5. Experimentation Using PCA with Rescaling Down Factor

As there are 10 subjects in the database and six images per subject in Subset1, the total number of images used as training images is 60. All these images are used to find the face-space in k -NNC. The mean image of a subject is cultivated using its six images of Subset1. Thus there are 10 mean images corresponding to 10 subjects of the database. The face-space in NMC is generated using these mean images. If we compare the dimension of face-space generated using PCA with the dimension of face-space used in correlation coefficients calculation, we see a large reduction in the dimension.

There is a slight reduction in the dimension of face-space used in eyes, nose and mouth region based in comparison with full-face image based. In eyes, nose and mouth region based method of correlation coefficient calculation, the dimension of face-space of training set is 5600×60 for k -NNC and this is 5600×10 for NMC. Whereas for full-face image based method, the dimension of face-space of training set is 12600×60 for k -NNC and this is 12600×10 for NMC. For principal component analysis (eigenface method), we have taken 50 principal components in our implementation of

this algorithm. Thus the dimension of face-space in PCA is 50×60 for k -NNC and this is 50×10 for NMC. So this dimension reduction results in much smaller computational complexity incurred in k -NNC and NMC used with PCA. This result is in the cost of increased no illumination compensation recognition error. With all these variations of different algorithms, our method of illumination normalization provides zero recognition error, thus the robustness of proposed technique is proved.

The dependency of recognition error on RDF is plotted in Figure 10. For RDF ranging from 11 to 89, the recognition error is zero for all test subsets when classification is done using k -NNC. As in the analysis done with full-face correlation coefficient with NMC, here also if RDF is above 13, the recognition error is zero. The mean value of RDF for further analysis is taken as mean of 11 and 89, which is 50 considering k -nearest neighbor classifier. In this analysis also, we have taken the value of C_{resc} as equal to 37.

4.6. Experimentation Using PCA with Number of Low Frequency DCT Coefficients

Figure 11 shows the effect of number of low-frequency DCT coefficients selected for re-scaling (C_{resc}) on recognition error. As we found with correlation coefficient based analysis, here also the recognition error decreases significantly after a small number of coefficients selected for re-scaling. The percentage recognition error is

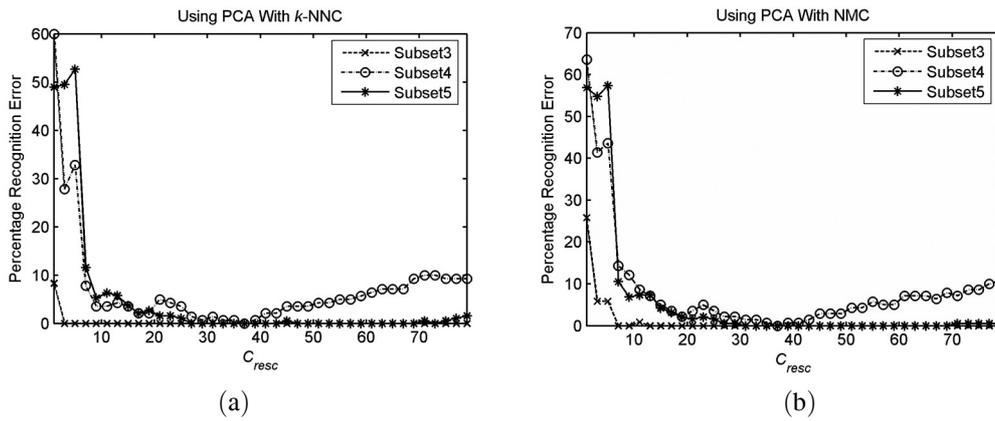


Figure 11. Performance variation using PCA with different value of C_{resc} for (a) k -NNC (b) NMC.

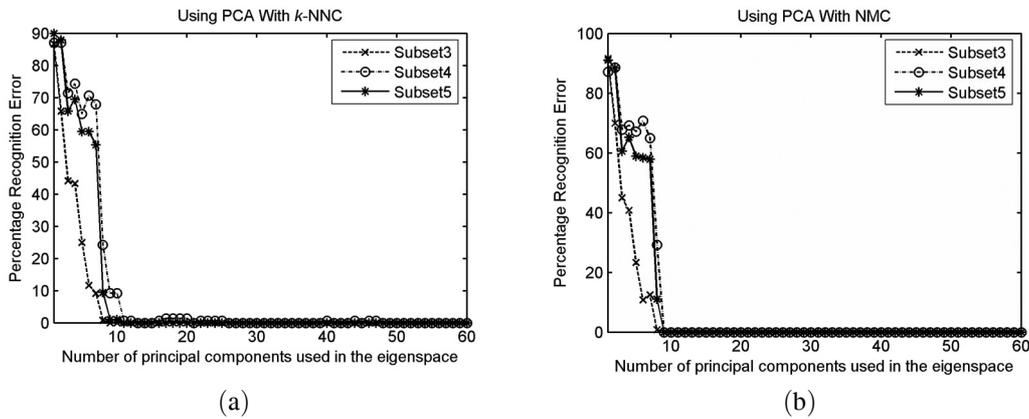


Figure 12. Performance variation using PCA with number of principal components for (a) k -NNC (b) NMC.

less than 10% for k -NNC when the value of C_{resc} is greater than 7 and the recognition error is zero for both classifiers for $C_{resc} = 37$. As aforementioned, the value of RDF is 50 in this analysis.

4.7. Experimentation Using PCA with Number of Principal Components

The performance variation with the dimension of eigenspace (number of principal components) is not yet studied in the analysis of illumination normalization. The graphs in Figure 12 present the analysis of this dependency. It is evident that the smaller number of principal components is sufficient to do accurate classification. The number of principal components is around 13 for k -NNC to do correct classification and

this value is even less than 10 in case of the nearest mean classifier. Thus for real-time applications, the reduced dimension of eigenspace can be useful as this reduces the computational complexity.

4.8. Discussion and Comparison with Other Approaches

The percentage recognition errors are summarized in Table 2. The recognition error with Subset2 is zero or almost negligible since this set contains face image with small illumination variations. So in all performance graphs and in Table2, the recognition error with Subset2 is not shown. Apart from achieving the accurate

Schemes	Percentage Recognition Error using k -NNC			Percentage Recognition Error using NMC		
	Subset3	Subset4	Subset5	Subset3	Subset4	Subset5
Without illumination normalization using correlation coefficient between eyes, nose and mouth region	0.83	31.43	31.58	12.50	48.57	47.89
Without illumination normalization using correlation coefficient between full-face image	8.3	60.00	49.47	27.5	62.14	57.37
Without illumination normalization using euclidean distance between PCA weight Vectors	8.3	60	48.95	25.83	63.57	56.84
With our approach of illumination normalization using any of the above mentioned distance metric	0	0	0	0	0	0

Table 2. Summary of our implementation results.

Schemes	Percentage Recognition Error		
	Subset3	Subset4	Subset5
Illumination Ratio Images [26]	3.3	18.6	Not available
Harmonic Images [11]	0.3	3.1	Not available
Linear Subspace [9]	0	15.0	Not available
Cones-Attached [9]	0	8.6	Not available
Cones-Cast [9]	0	0	Not available
Gradient Angle [28]	0	1.4	Not available
9PL [29]	0	2.8	Not available
Quotient Illumination Relighting [27]	0	9.4	17.5
Our Method	0	0	0

Table 3. Performance comparison of different schemes

face recognition using our method of illumination normalization, we have tested the performance of different methods of face recognition with two variants of nearest neighbor classifier, k -NNC and NMC. If we compare recognition errors, the without illumination-normalization correlation coefficient using eyes, nose and mouth region based, performs better in comparison to other two methods. Thus local feature based approach performs better in comparison to appearance (global) based approach. It is also evident from Table 2 and different performance graphs that the robustness of NMC is proven in presence of different parameter variation, whereas the k -NNC gives better performance in no illumination normalization.

In Table 3 the recognition errors on Yale Face Database B using other existing methods deal-

ing with illumination normalization are presented along with the proposed approach. The results of existing methods are cited in the reference directly from their and other papers. Most of the existing methods are not finding recognition error for Subset5 as this set contains images corresponding to extreme illumination variations. The last two existing methods cited above our method in Table 3 have larger recognition error compared to the error for Subset4. Using our method, the recognition errors for all subsets are zero. If we compare our implementation results for Subset4 and Subset5 without illumination normalization (Table2), we find that the recognition error for Subset5 is less than that for Subset4. This became possible as all images are histogram equalized before applying DCT and the effect of HE is more for Subset5 compared

to Subset4. As a result the images of Subset5 become closer to training image(s).

5. Conclusions

This paper proposes a novel approach for illumination normalization under varying lighting conditions. The illumination normalization is done in DCT domain using low-frequency DCT coefficients re-scaling. The experimental results show that proper re-scaling down of sufficient number of low-frequency DCT coefficients can result in 100% accurate face recognition. We have also tested the performance of different face recognition methods with our illumination normalization approach. The experimental results show that local feature based approach (correlation coefficient based on eyes, nose and mouth region) performs better in comparison to appearance based approach (correlation coefficient based on full-face image and PCA). It is also evident that the nearest mean classifier is more robust in comparison to k -NNC for parameter change. In the present research work, only frontal images with varying illumination are considered. In our future work, we will focus on the illumination normalization in the presence of pose and expression changes.

6. Acknowledgment

The authors would like to express their sincere gratitude to Yale University for the use of Yale Face Database B.

References

- [1] W. ZHAO, R. CHELLAPPA, P. J. PHILLIPS AND A. ROSENFELD, Face Recognition: A Literature Survey, *ACM Computing Surveys* 35 (4) (2003) pp. 399–458.
- [2] R. CHELLAPPA, C. L. WILSON AND S. SIROHEY, Human and machine recognition of faces: a survey. *Proceedings of IEEE* 83 (5) (1995) pp. 705–740.
- [3] P. N. BELHUMEUR AND D. J. KRIEGMAN, What is the set of images of an object under all possible illumination conditions. *International Journal of Computer Vision* 28 (3) (1998) pp. 245–260.
- [4] S. Z. LI AND A. K. JAIN, *Handbook of Face Recognition*, Springer, 2005.
- [5] X. ZOU, J. KITTLER AND K. MESSER, Illumination Invariant Face Recognition: A Survey, in: *Proceedings of 1st IEEE International Conference on Biometric: Theory, Application and Systems* (2007), pp. 1–8.
- [6] Z. M. HAPED AND M. D. LEVINE, Face Recognition Using the Discrete Cosine Transform. *International Journal of Computer Vision* 43 (3) (2001) pp. 167–188.
- [7] P. N. BELHUMEUR, J. P. HESPANHA AND D. J. KRIEGMAN, Eigenfaces versus Fisherfaces: recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* 19 (7) (1997) pp. 711–720.
- [8] A. SHASHUA AND T. RIKLIN-RAVIV, The quotient image: class-based re-rendering and recognition with varying illuminations. *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (2) (2001) pp. 129–139.
- [9] A. S. GEORGHIADES, P. N. BELHUMEUR AND D. W. JACOBS, From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (6) (2001) pp. 630–660.
- [10] R. BASRI AND D. W. JACOBS, Lambertian reflectance and linear subspaces. *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (2) (2003) pp. 218–233.
- [11] L. ZHANG AND D. SAMARAS, Face recognition under variable lighting using harmonic image exemplars, in: *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 1, 2003, pp. 19–25.
- [12] S. S. AGAIAN, B. SILVER AND K. A. PANETTA, Transform Coefficient Histogram-Based Image Enhancement Algorithms Using Contrast Entropy, *IEEE Trans. on Image Processing* 16 (3) (2007) pp. 741–758.
- [13] R. C. GONZALEZ AND R. E. WOODS, *Digital Image Processing*, Pearson Education, 2006.
- [14] A. K. JAIN, *Fundamentals of Digital Image Processing*, Pearson Education, 2006.
- [15] T. RONNY, L. WANQUAN AND V. SVETHA, Application of the DCT Energy Histogram for Face Recognition, in: *Proceedings of 2nd International Conference on Information Technology for Application (ICITA 2004)*, Jan. 2004, pp. 305–310.
- [16] R. BRUNELLI AND T. POGGIO, Face Recognition: Features vs. Templates, *IEEE Trans. Pattern Anal. Mach. Intell.* 15 (10) (1993) pp. 1042–1053.
- [17] F. ZHAO, Q. HUANG AND W. GAO, Image Matching by Normalized Cross-Correlation, in: *Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006)* vol. 2, May 2006, pp. 729–732.
- [18] MATLAB REFERENCE MANUAL, <http://www.mathworks.com/access/helpdesk/help/toolbox/images>, 2004.

- [19] M. TURK AND A. PENTLAND, Eigenfaces for Recognition, *J. Cognitive Neuroscience*, 3 (1) (1991) pp. 71–96.
- [20] K. DELAC, M. GRGIC AND S. GRGIC, Independent Comparative Study of PCA, ICA, and LDA on the FERET Data Set, *International Journal of Imaging Systems and Technology*, 15 (5) (2006) pp. 252–260.
- [21] T. M. COVER AND P. E. HART, Nearest Neighbor Pattern Classification, *IEEE Trans. Information Theory* 13 (1) (1967) pp. 21–27.
- [22] A. DJOUADI AND E. BOUKTACHE, A Fast Algorithm for the Nearest-Neighbor Classifier, *IEEE Trans. Pattern Anal. Mach. Intell.*, 19 (3) (1997) pp. 277–281.
- [23] C. J. VEENMAN AND M. J. T. REINDERS, The Nearest Subclass Classifier: A Compromise between the Nearest Mean and Nearest Neighbor Classifier, *IEEE Trans. Pattern Anal. Mach. Intell.*, 23 (9) (2005) pp. 1417–1429.
- [24] Z. LOU AND Z. JIN, Novel Adaptive Nearest Neighbor Classifiers Based On Hit-Distance, in: *Proceedings of 18th International Conference on Pattern Recognition (ICPR 2006)*, vol. 3, 2006, pp. 87–90.
- [25] H. SAMET, K-Nearest Neighbor Finding Using MaxNearestDist, *IEEE Trans. Pattern Anal. Mach. Intell.* 30 (2) (2008) pp. 243–252.
- [26] J. ZHAO, Y. SU, D. WANG, AND S. LUO, Illumination ratio image: synthesizing and recognition with varying illuminations, *Pattern Recognition Letter*, 24 (2003) pp. 2703–2710.
- [27] S. SHAN, W. GAO, B. CAO AND D. ZHAO, Illumination normalization for robust face recognition against varying lighting conditions, in: *Proceedings of IEEE Workshop on AMFG*, 2003, pp. 157–164.
- [28] H. F. CHEN, P. N. BELHUMEUR AND D. J. KRIEGMAN, In search of illumination invariants, in: *Proceedings of IEEE International Conference Computer Vision and Pattern Recognition*, vol. 1, 2000, pp. 13–15.
- [29] K.-C. LEE, J. HO AND D. J. KRIEGMAN, Acquiring linear subspaces for face recognition under variable lighting, *IEEE Trans. Pattern Anal. Mach. Intell.*, 27 (5) (2005) pp. 684–698.
- [30] YALE FACE DATABASE B, <http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html>, 2001.
- [31] P. J. PHILLIPS, H. MOON, S. A. RIJVI AND P. J. RAUSS, “The FERET Evaluation Methodology for Face-Recognition Algorithms,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 22 (10) (2000) pp. 1090–1104.
- [32] H. BISCHOF, H. WILDENAUER AND A. LEONARDIS, “Illumination insensitive recognition using eigenspaces,” *Computer Vision and Image Understanding*, 95 (1) (2004) pp. 86–104.

Received: January, 2009
 Revised: October, 2009
 Accepted: October, 2009

Contact addresses:

Virendra P. Vishwakarma
 Amity School of Engineering and Technology
 (Guru Gobind Singh Indraprastha University)
 580, Delhi Palam Vihar Road, Bijwasan,
 New Delhi-110061, India
 e-mail: vpvishwakarma@aset.amity.edu,
 virendravishwa@rediffmail.com

Sujata Pandey
 Amity School of Engineering and Technology
 (Guru Gobind Singh Indraprastha University)
 580, Delhi Palam Vihar Road, Bijwasan,
 New Delhi-110061, India
 e-mail: spandey@amity.edu

M. N. Gupta
 Amity School of Engineering and Technology
 (Guru Gobind Singh Indraprastha University)
 580, Delhi Palam Vihar Road, Bijwasan,
 New Delhi-110061, India
 e-mail: mngupta@gmail.com

VIRENDRA P. VISHWAKARMA received his B.Tech. degree in Electrical Engineering from H.B.T.I. Kanpur, Uttar Pradesh, India, in 1994; M.E. in Computer Science & Engineering from M.N.N.I.T Allahabad, Uttar Pradesh, India, in 2002. He is presently working towards his Ph.D. degree from School of Engineering and Technology, Guru Gobind Singh Indraprastha University, New Delhi, India. From 1996 to 2000, he worked for Research Design and Standards Organization, Ministry of Railways, India, where he was involved in many research projects. He also worked in Research & Development division of Himachal Futuristic Communication Limited, Gurgaon, India for 7 months. He joined the Department of Computer Science & Engineering at Amity School of Engineering & Technology, New Delhi, India in 2002 and presently he is working as Associate Professor. His present research interests include digital image processing, pattern recognition (face recognition) and soft computing. He is member of IACSIT, ITHEA and ACEEE.

SUJATA PANDEY received Masters degree in Electronics (VLSI) from Kurukshetra University in 1994 and Ph.D. degree from Department of Electronics, University of Delhi South Campus, India in Microelectronics in 1999. She joined Semiconductor Devices Research Laboratory, University of Delhi in 1996 under the project of CSIR, Ministry of Science and Technology, Govt. of India. She joined the Department of Electronics & Communication Engineering, Amity School of Engineering & Technology, New Delhi, India in 2002 as Assistant Professor. She has published more than 80 research papers in International/National Journals/Conferences. Her current research interests include modeling and characterization of HEMTs, SOI devices integrated circuit design and digital image processing. She is member of IEEE and Electron Device Society.

M. N. GUPTA received Ph.D. degree from I.I.T. Delhi, India, in Computer Science and Engineering in 1989. He has run many research projects and has published many research papers in reputed International Journals/Conferences. He has more than 30 years of academic experience in various organizations such as IIT Delhi, University of Delhi, Amity School of Engineering and Technology India. He is currently professor and Head of Computer Science & Engineering and Information Technology Departments at Amity School of Engineering and Technology, New Delhi, India. His current research interests include digital image processing and distributed computing.
