Research Article

Data Fusion Boosted Face Recognition Based on Probability Distribution Functions in Different Colour Channels

Hasan Demirel (EURASIP Member) and Gholamreza Anbarjafari

Department of Electrical and Electronic Engineering, Eastern Mediterranean University, Gazimağusa, KKTC, 10 Mersin, Turkey

Correspondence should be addressed to Hasan Demirel, hasan.demirel@emu.edu.tr

A new and high performance face recognition system based on combining the decision obtained from the probability distribution functions (PDFs) of pixels in different colour channels is proposed. The PDFs of the equalized and segmented face images are used as statistical feature vectors for the recognition of faces by minimizing the Kullback-Leibler Divergence (KLD) between the PDF of a given face and the PDFs of faces in the database. Many data fusion techniques such as median rule, sum rule, max rule, product rule, and majority voting and also feature vector fusion as a source fusion technique have been employed to improve the recognition performance. The proposed system has been tested on the FERET, the Head Pose, the Essex University, and the Georgia Tech University face databases. The superiority of the proposed system has been shown by comparing it with the state-of-art face recognition systems.

1. Introduction

The earliest work in computer recognition of faces was reported by Bledsoe [1], where manually located feature points are used. Statistical face recognition systems such as principal component analysis- (PCA-) based eigenfaces introduced by Turk and Pentland [2] attracted a lot of attention. Belhumeur et al. [3] introduced the fisherfaces method which is based on linear discriminant analysis (LDA).

Many of these methods are based on greyscale images; however colour images are increasingly being used since they add additional biometric information for face recognition [4]. Colour PDFs of a face image can be considered as the signature of the face, which can be used to represent the face image in a low-dimensional space. Images with small changes in translation, rotation, and illumination still possess high correlation in their corresponding PDFs, which prompts the idea of using PDFs for face recognition.

PDF of an image is a normalized version of an image histogram. Hence the published face recognition papers using histograms indirectly use PDFs for recognition, there is some published work on application of histograms for the detection of objects [5]. However, there are few publications on application of histogram or PDF-based methods in face recognition: Yoo and Oh used chromatic histograms of faces [6]. Ahonen et al. [7] and Rodriguez and Marcel [8] divided a face into several blocks and extracted the Local Binary Pattern (LBP) feature histograms from each block and concatenated into a single global feature histogram to represent the face image; the face was recognized by a simple distance based grey-level histogram matching. Demirel and Anbarjafari [9] introduced high performance pose invariant face recognition system based on greyscale histogram of faces, where the cross-correlation coefficient between the histogram of the query image and the histograms of the training images was used as a similarity measure.

Face segmentation is one of the important preprocessing phases of face recognition. There are several methods for this task such as skin tone-based face detection for face segmentation. Skin is a widely used feature in human image processing with a range of applications [10]. Human skin can be detected by identifying the presence of skin colour pixels. Many methods have been proposed for achieving this. Chai and Ngan [11] modelled the skin colour in *YCbCr* colour space. One of the recent methods for face detection is proposed by Nilsson et al. [12] which is using local Successive
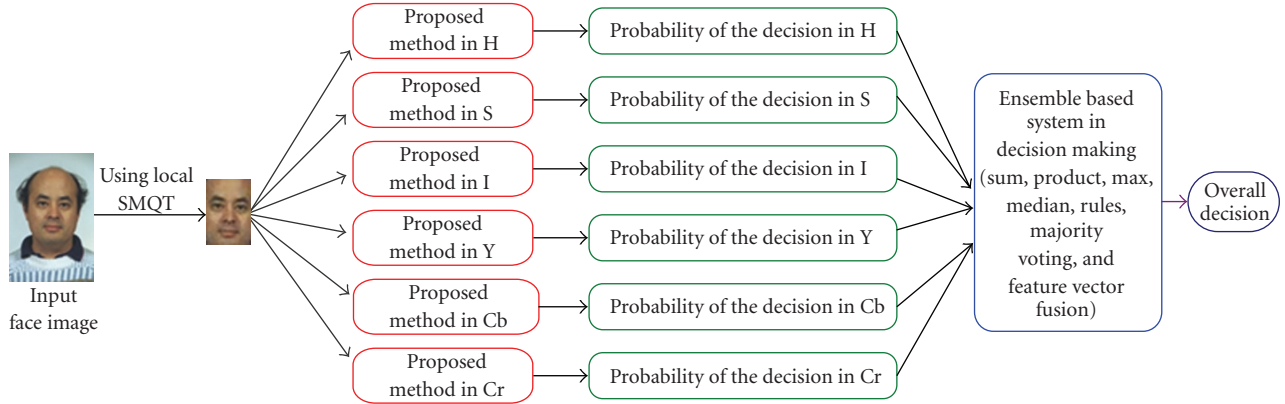
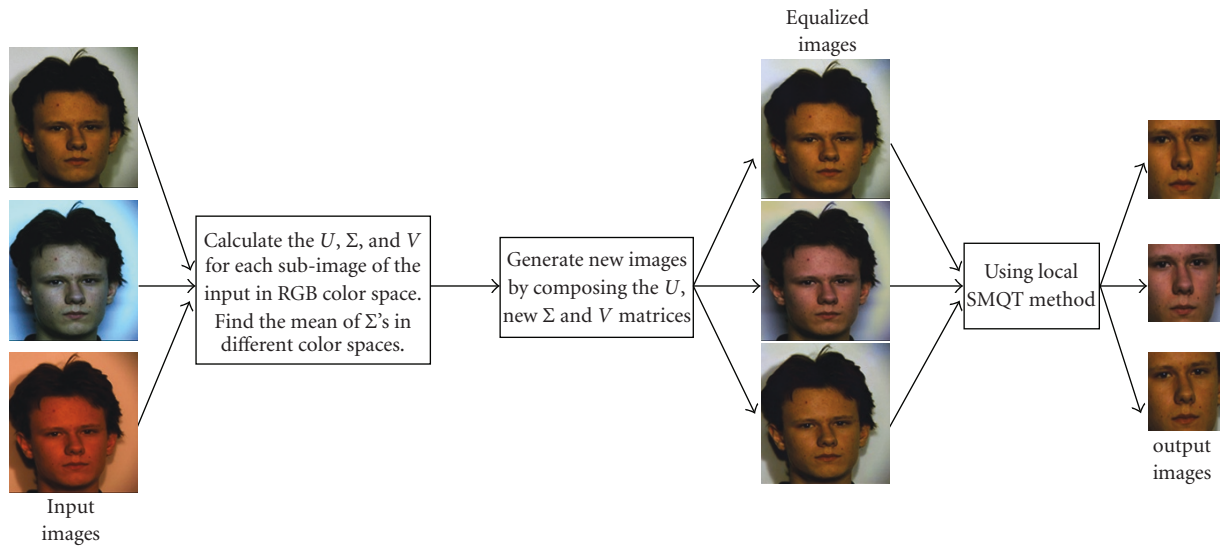FIGURE 1: Different phases of the proposed system.

FIGURE 2: The algorithm, with a sample image with different illumination from Oulu face database, of pre-processing of the face images to obtain a segmented face from the input face image.

TABLE 1: The entropy of colour images in different colour channels compared with the greyscale images.

Database	The average entropy of the images (bits/pixel)		
	HSI	*YCbCr*	Greyscale
FERET	19.2907	16.3607	7.1914
Head Pose	15.9434	12.3173	6.7582
Essex Uni.	21.2082	17.3158	7.0991
Georgia Tech	20.8015	16.6226	6.9278

Mean Quantization Transform (SMQT) technique. Local SMQT is robust for illumination changes, and the Receiver Operation Characteristics of the method are reported to be very successful for the segmentation of faces.

In the present paper, the local SMQT algorithm has been adopted for face detection and cropping in the pre-processing stage. Colour PDFs in *HSI* and *YCbCr* colour spaces of the isolated face images are used as the face des-

criptors. Face recognition is achieved using the Kullback-Leibler Divergence (KLD) between the PDF of the input face and the PDFs of the faces in the training set. Different data and source fusion methods have been used to combine the decision of the different colour channels to increase the recognition performance. In order to reduce the effect of the illumination, the singular value decomposition-based image equalization has been used. **Figure 1** illustrates the phases of the proposed system which combines the decisions of the classifiers in different colour channels for improved recognition performance.

The system has been tested on the Head Pose (HP) [13], FERET [14], Essex University [15] and the Georgia Tech University [16] face databases where the faces have more varying background and illumination than pose changes.

2. Preprocessing of Face Images

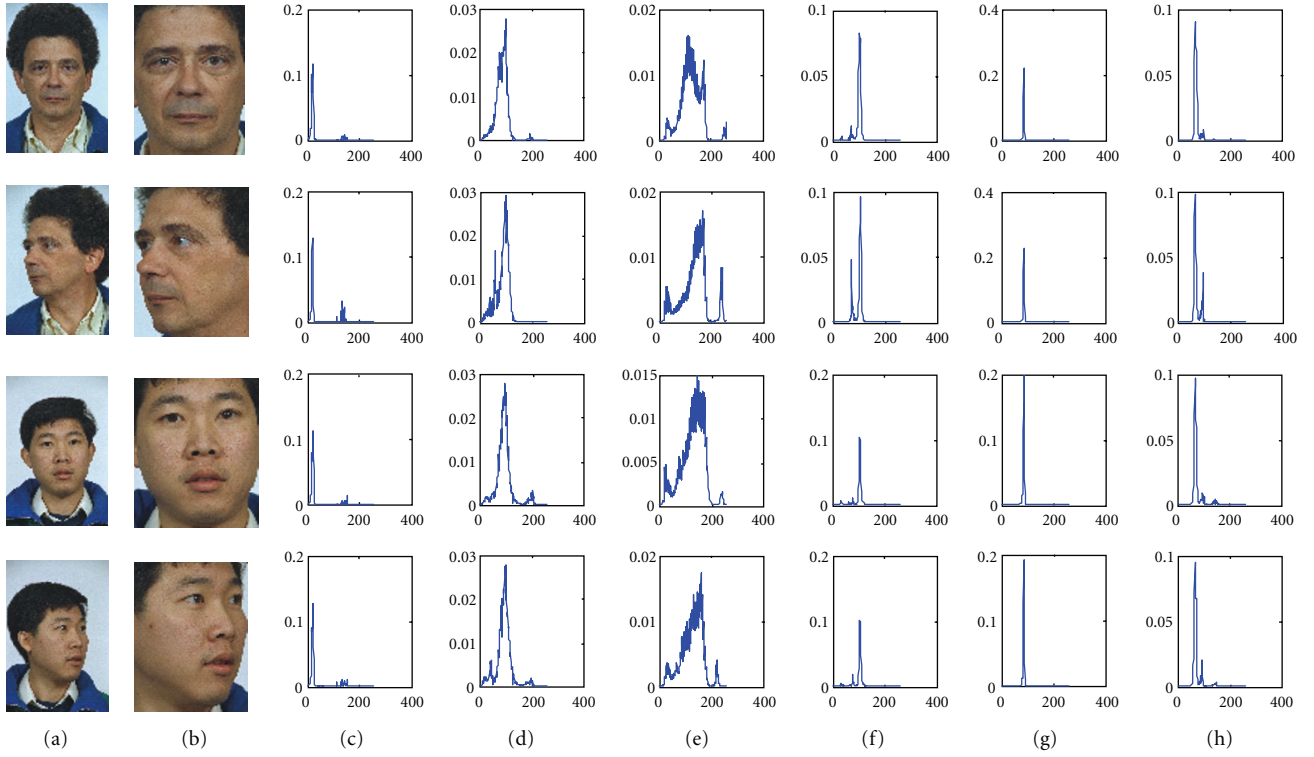There are several approaches used to eliminate the illumination problem of the colour images [17]. One of the most

FIGURE 3: Two subjects from FERET database with 2 different poses (a), their segmented faces (b) and their PDFs in H (c), S (d), I (e), Y (f), Cb (g), and Cr (h) colour channels respectively.

TABLE 2: Performance of the proposed PDF-based face recognition system in H, S, I, Y, Cb, and Cr colour channels of the FERET, HP, Essex, and Georgia Tech University face databases.

Database	No of training per subject	Colour channels					
		H	S	I	Y	Cb	Cr
FERET	1	77.64	60.16	49.24	56.89	57.42	67.40
	2	84.75	68.03	58.43	65.60	66.50	73.40
	3	91.63	76.20	67.89	74.63	75.46	81.23
	4	93.67	81.70	73.03	79.00	81.00	84.27
	5	95.08	83.52	78.84	84.20	85.00	88.32
HP	1	66.96	62.52	48.07	54.81	62.44	75.19
	2	83.17	76.17	70.25	78.00	77.75	86.08
	3	85.81	75.43	75.81	84.10	78.95	90.19
	4	89.78	83.89	80.67	88.00	85.56	92.78
	5	88.80	87.87	87.07	93.87	84.53	94.13
Essex Uni.	1	73.27	61.05	81.21	90.19	73.90	76.66
	2	83.37	68.03	85.35	94.79	82.53	85.26
	3	87.22	70.17	87.45	96.84	86.62	88.05
	4	90.13	72.07	88.64	97.09	88.26	90.03
	5	90.72	74.57	89.12	97.85	90.01	91.31
Georgia Tech Uni.	1	67.13	65.13	66.11	67.07	64.69	66.73
	2	84.05	81.18	82.63	81.58	82.65	83.28
	3	89.74	87.71	89.11	89.26	87.77	87.4571
	4	92.63	91.20	91.73	91.77	91.87	91.07
	5	94.60	93.08	93.52	93.08	93.88	92.48

TABLE 3: Performance of the PCA-based system in *H, S, I, Y, Cb,* and *Cr* colour channels of different face databases.

Database	No of training per subject	Colour channels					
		H	*S*	*I*	*Y*	*Cb*	*Cr*
FERET	1	36.89	48.67	44.00	47.33	49.78	49.11
	2	41.50	54.75	52.00	52.50	58.25	57.75
	3	52.86	62.86	58.29	56.57	67.71	64.00
	4	58.00	69.00	66.17	66.00	73.67	70.33
	5	62.40	74.80	68.80	72.80	77.60	74.80
HP	1	12.59	17.78	20.74	20.74	20.00	18.52
	2	24.17	38.33	41.67	43.33	38.33	31.67
	3	30.48	57.14	56.19	59.05	53.33	45.71
	4	32.22	58.89	58.89	62.22	55.56	50.00
	5	38.67	62.67	66.67	69.33	65.33	56.00
Essex Uni.	1	74.04	89.07	93.16	92.80	91.29	92.44
	2	83.60	93.00	94.20	94.20	94.70	92.40
	3	85.14	94.06	94.06	94.29	95.20	93.14
	4	87.20	94.40	93.87	94.27	95.47	93.07
	5	88.96	96.16	94.88	95.68	96.16	94.08
Georgia Tech Uni.	1	46.44	62.44	52.89	54.22	57.33	54.89
	2	51.75	68.75	59.00	59.25	67.00	58.50
	3	51.43	67.71	58.86	58.86	65.71	59.14
	4	50.00	65.33	58.67	60.00	63.33	57.67
	5	49.60	61.20	60.80	60.80	65.60	56.40

TABLE 4: Performance of different decision making techniques for the proposed face recognition system.

	No of training image per subject	Sum rule	Median rule	Min rule	Product rule	Majority voting	Feature vector Fusion
Head pose	1	83.85	84.74	74.74	84.22	81.04	81.48
	2	96.42	97.00	88.17	97.33	92.17	87.50
	3	96.76	96.19	90.95	96.86	93.43	96.19
	4	96.67	97.00	91.67	97.11	95.78	97.33
	5	97.33	98.53	91.47	96.27	97.33	97.78
FERET	1	76.89	76.80	66.87	75.60	75.22	82.89
	2	87.63	88.10	79.98	48.95	86.03	87.00
	3	89.97	90.26	82.6	14.83	88.54	96.57
	4	93.80	93.50	87.07	4.83	92.20	98.80
	5	95.16	95.44	89.84	4.00	94.20	99.33
Essex	1	94.53	93.82	81.71	16.58	92.45	95.33
	2	97.03	96.23	87.78	0.67	95.51	97.58
	3	98.08	97.80	90.37	0.67	96.55	97.81
	4	98.49	97.98	91.83	0.67	96.88	97.33
	5	98.84	98.39	92.87	0.67	97.41	97.73
Georgia	1	69.24	69.24	68.51	68.98	69.04	73. 20
	2	86.35	86.45	85.05	64.25	85.65	78.67
	3	90.71	90.83	89.91	23.97	91.46	74.78
	4	95.33	95.20	93.17	6.80	94.63	72.54
	5	95.96	96.04	95.48	3.20	95.84	75.29

frequently used and simplest methods is to equalize the colour image in RGB colour space by using histogram equalization (HE) in each colour channel separately. Previously we proposed singular value equalization (SVE) technique which is based on singular value decomposition (SVD) to equalize an image [18, 19]. In general, for any intensity image matrix $\Xi_A, A = \{R, G, B\}$, SVD can be written as

$$\Xi_A = U_A \Sigma_A V_A^T, \quad A = \{R, G, B\}, \tag{1}$$

where U_A and V_A are orthogonal square matrices (hanger and aligner matrices), and Σ_A matrix contains the sorted singular values on its main diagonal (stretcher matrix). As reported in [20], Σ_A represents the intensity information of a given image intensity matrix. If an image is a low contrast image this problem can be corrected to replace the Σ_A of the image with another singular matrix obtained from a normal image with no contrast problem. Any pixel of an image can be considered as a random value with distribution function of Ψ. According to the central limit theorem (CLT), the normalized sum of a sequence of random variables tends to have a standard normal distribution with mean 0 and standard deviation 1, which can be formulated as follows:

$$\lim_{n \to \infty} P(\mathbb{Z}_n \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx,$$

$$\text{where } \mathbb{Z}_n = \frac{S_n - E(S_n)}{\sqrt{\text{var}(S_n)}}, \quad S_n = \sum_{i=1}^{n} X_i. \tag{2}$$

Hence a normalized image with no intensity distortion (i.e., no external condition forces the pixel value to be close to a specific value, thus the distribution of each pixel is identical) has a normal distribution with mean of 0 and variance of 1. Such a synthetic matrix with the same size of the original image can easily be obtained by generating random pixel values with normal distribution with mean of 0 and variance of 1.

Then the ratio of the largest singular value of the generated normalized matrix over a normalized image can be calculated according to

$$\xi_A = \frac{\max\left(\Sigma_{N(\mu=0,\sigma=1)}\right)}{\max(\Sigma_A)}, \quad A = \{R, G, B\}, \tag{3}$$

where $\Sigma_{N(\mu=0,\sigma=1)}$ is the singular value matrix of the synthetic intensity matrix. This coefficient can be used to regenerate a new singular value matrix which is actually an equalized intensity matrix of the image generated by

$$\Xi_{\text{equalized}_A} = U_A(\xi_A \Sigma_A) V_A^T, \quad A = \{R, G, B\}, \tag{4}$$

where $\Xi_{\text{equalized}_A}$ is representing the equalized image in A-colour channel.

As (4) states, the equalized image is just a multiplication of ξ_A with the original image. From the computational complexity point of view singular value decomposition of a matrix is an expensive process which takes quite significant amount of time to calculate the orthogonal matrices of U_A and V_A while they are not being used in the equalization

process. Hence, finding a cheaper method to obtain ξ can be an improvement to the technique. Recall

$$\|A\| = \sqrt{\lambda_{\max}}, \tag{5}$$

where $\lambda_{\max}$ is the maximum eigenvalue of $A^T A$. By using SVD,

$$A = U\Sigma V^T \rightarrow A^T A = V\Sigma^2 V^T. \tag{6}$$

This follows that the eigenvalues of $A^T A$ are the square of elements of the main diagonal of Σ, and that the eigenvector of $A^T A$ is V. Because Σ is in the form of

$$\Sigma = \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \lambda_k \cdots \end{bmatrix}_{m \times n}, \tag{7}$$

$$\lambda_1 > \lambda_2 > \cdots > \lambda_k, \quad k = \min(m, n)$$

where λ_i is the ith eigenvalue of A. Thus,

$$\|A\| = \lambda_1. \tag{8}$$

The 2-norm of a matrix is equal to the largest singular value of the matrix. Therefore ξ_A can be easily obtained from

$$\xi_A = \frac{\left\|\Xi_{N(\mu=0,\sigma=1)}\right\|}{\|\Xi_A\|}, \quad A = \{R, G, B\}, \tag{9}$$

where $\Xi_{N(\mu=0,\sigma=1)}$ is a random matrix with mean of 0 and variance of 1, and Ξ_A is the intensity image in R, G, or B. Hence the equalized image can be obtained by

$$\Xi_{\text{equalized}_A} = \xi_A \Xi_A = \frac{\left\|\Xi_{N(\mu=0,\sigma=1)}\right\|}{\|\Xi_A\|} \Xi_A, \quad A = \{R, G, B\}, \tag{10}$$

which shows there is no need to use singular value decomposition of intensity matrices. This procedure eases the equalization step. Note that, Ξ_A is a normalized image with intensity values between 0 and 1. After generation of Ξ_N, it is normalized such that the values are between 0 and 1.

This task which is actually equalizing the images of a face subject will eliminate the illumination problem. Then, this new image can be used as an input for the face detector prepared by Nilsson [21] in order to segment the face region and eliminate the undesired background.

The local successive mean quantization transform (SMQT) can be explained as follows. The SMQT can be considered as an adjustable tradeoff between the number of quantization levels in the result and the computational load [22]. Local is defined to be the division of an image into blocks with a predefined size. Let x be a pixel of local D, and let us have the SMQT transform as follows:

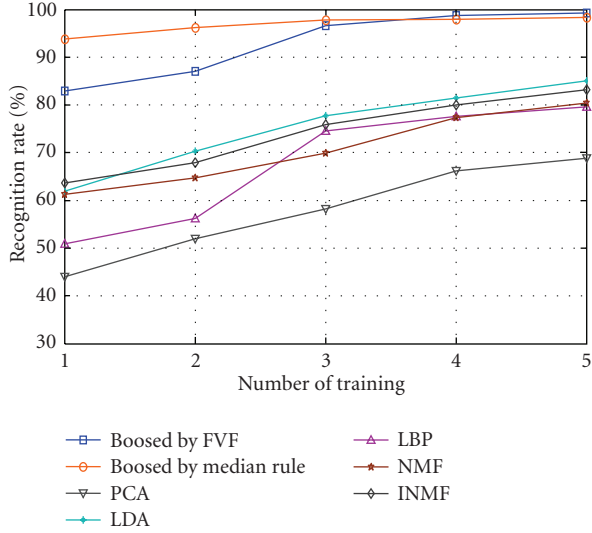$$\text{SMQT}_L : D(x) \rightarrow M(x), \tag{11}$$

FIGURE 4: Recognition rate (%) vs. number of training faces for the FERET face database, using proposed FVF and median rule based systems compared with PCA , LDA, LBP, NMF, and INMF.

where $M(x)$ is a new set of values which are insensitive to gain and bias [22]. These two properties are desired for the formation of the intensity image which is a product of reflection and illumination. A common approach to separate the reflection and illumination is based on this assumption that illumination is spatially smooth so that it can be taken as a constant in a local area. Therefore each local pattern with similar structure will yield the similar SMQT features for a specified level, L. The spare network of winnows (SNoWs) learning architecture is also employed in order to create a look-up table for classification. As Nilsson et al. proposed in [22], in order to scan an image for faces, a patch of 32×32 pixels is used and also the image is downscaled and resized with a scale factor to enable the detection of faces with different sizes. The choice of the local area and the level of the SMQT are vital for successful practical operation. The level of the transform is also important in order to control the information gained from each feature. As reported in [22] the 3×3 local area and level $L = 1$ are used to be a proper balance for the classifier. The face and nonface tables are trained in order to create the split up SNoW classifier. Overlapped detections are disregarded using geometrical locations and classification score. Hence given two detections overlapping each other, the detection with the highest classification score is kept and the other one is removed. This operation is repeated until no overlapping detection is found.

The segmented face images are used for the generation of PDFs in H, S, I, Y, Cb, and Cr colour channels in HSI and $YCbCr$ colour spaces. If there is no face in the image, then there will be no output from the face detector software, so it means the probability of having a random noise which has the same colour distribution of a face but with different shape is zero, which makes the proposed method reliable. The proposed equalization has been tested on the Oulu face database [23] as well as the FERET, the HP,

the Essex University, and the Georgia Tech University face databases. Figure 2 shows the general required steps of the preprocessing phase of the proposed system.

3. Colour Images versus Greyscale Images

Usually many face recognition systems use greyscale face images. From the information point of view a colour image has more information than a greyscale image. So we propose not to lose the available amount of information by converting a colour image into a greyscale image. In order to compare the amount of the information in a colour and greyscale images, the entropy of an image can be used, which can be calculated by

$$H = -\sum_{\xi=0}^{255} P(\zeta)\log_2(P(\zeta)), \qquad (12)$$

where H measures the information of the image. The average amount of information measured by using 2650 face images of the FERET, HP, Essex University, and Georgia Tech University face databases is shown in Table 1. The entropy values indicate that there is significant amount of information in different colour channels which should not be simply ignored by only considering the greyscale image.

4. PDF-Based Face Recognition

The PDF of an image is a statistical description of the distribution in terms of occurrence probabilities of pixel intensities, which can be considered as a feature vector representing the image in a lower-dimensional space [18]. In a general mathematical sense, an image PDF is simply a mapping η_i representing the probability of the pixel intensity levels that fall into various disjoint intervals, known as bins. The bin size determines the size of the PDF vector. In this work the bin size is assumed to be 256. Given a monochrome image, PDF η_j meet the following conditions, where N is the total number of pixels in an image:

$$N = \sum_{j=0}^{255} \eta_j. \qquad (13)$$

Then, PDF feature vector, H, is defined by

$$H = [p_0, p_1, \dots, p_{255}], \qquad p_\iota = \frac{\eta_\iota}{N}, \quad \iota = 0, \dots, 255, \qquad (14)$$

where η_i is the intensity value of a pixel in a colour channel, and N is total number of pixels in an intensity image. Kullback-Leibler Divergence can be used to measure the distance between the PDF of two images, although in general it is not a distance metric. Kullback-Leibler Divergence is sometimes referred as Kullback-Leibler Distance (KLD) as well [24]. Given two PDF vectors p and q the KLD, κ, is defined as

$$\kappa_i(q, p_j) = \sum_j q_j \log\left(\frac{q_j}{p_{ij}}\right), \qquad (15)$$

$$j = 0, 1, 2, \dots, \beta - 1, \quad i = 1, \dots, M,$$

TABLE 5: Performance of different decision making techniques for the PCA-based face recognition system.

	No of training image per subject	Sum rule	Median rule	Min rule	Product rule	Majority voting
	1	23.70	22.22	17.78	22.96	22.22
	2	41.67	42.50	39.17	45.00	42.50
Head pose	3	62.86	60.00	56.19	67.62	58.10
	4	66.67	63.33	62.22	65.56	61.11
	5	66.67	68.00	69.33	69.33	68.00
	1	63.11	56.67	41.56	63.78	56.89
	2	67.50	64.00	47.75	69.50	62.75
FERET	3	72.86	68.00	57.14	74.00	65.71
	4	80.33	76.67	60.33	80.67	74.00
	5	84.40	80.80	65.60	83.60	77.60
	1	97.69	97.51	95.38	97.96	96.27
	2	97.60	97.10	95.60	97.70	96.40
Essex	3	97.49	97.37	95.66	97.73	96.80
	4	97.33	96.93	95.73	97.47	97.07
	5	98.24	97.92	97.12	98.24	98.24
	1	66.22	64.00	53.11	66.44	61.78
	2	72.50	71.50	57.50	73.25	69.75
Georgia	3	74.00	72.86	57.71	74.29	70.00
	4	74.00	74.00	55.67	74.33	69.00
	5	72.00	70.40	53.20	72.40	68.40

TABLE 6: Performance of the proposed face recognition system using FVF, Median Rule, PCA, LDA, LBP, NMF, and INMF based face recognition system for the FERET face databases.

# of training images	1	2	3	4	5
FVF	82.89	87.00	96.57	98.80	99.33
MEDIAN RULE	93.82	96.23	97.80	97.98	98.39
PCA	44.00	52.00	58.29	66.17	68.80
LDA	61.98	70.33	77.78	81.43	85.00
LBP	50.89	56.25	74.57	77.67	79.60
NMF	61.33	64.67	69.89	77.35	80.37
INMF	63.65	67.87	75.83	80.07	83.20

where β is the number of bins, and M is the number of images in the training set. In order to avoid the three undefined possibilities: division by zero in $\log(q_j/p_{ij})$ where $p_{ij} = 0$, or $\log(0)$ where $q_j = 0$, or both situation together, we have modified the formula into the following form:

$$\kappa_i\left(q, p_j\right) = \sum_j q_j \log\left(\frac{q_j + \delta}{p_{ij} + \delta}\right),$$

$$j = 0, 1, 2, \dots, \beta - 1, \quad i = 1, \dots, M, \qquad (16)$$

where $\delta \ll 1/\beta$, for example, $\delta = 10^{-7}$. One should note that for an image the $p_{ij}, q_j \in \mathbb{Z}^+$, that is, their minimum value

is zero and the maximum value can be the number of pixels in an image. Then, a given query face image, the PDF of the query image q can be used to calculate the KLD between q and PDFs of the images in the training samples as follows:

$$\chi_r = \min\left(\kappa_i\left(q, p_j\right)\right), \quad i = 1, \dots, M. \qquad (17)$$

Here, χ_r is the minimum KLD reflecting the similarity of the rth image in the training set and the query face. The image with the lowest KLD distance from the training face images is declared to be the identified image in the set. Figure 3 shows two subjects with two different poses and their segmented faces from the FERET face database which is well known in terms of pose changes and also the images have different backgrounds with slight illumination variation. The intensity of each image has been equalized by using SVE to minimize the illumination effect.

The colour PDFs used in the proposed system are generated only from the segmented face, and hence the effect of background regions is eliminated. The performance of the proposed system is tested on the FERET, the HP, Essex University, and Georgia Tech University face databases with changing poses, background, and illumination, respectively. The details of these databases are given in Results and Discussions section. The faces in those datasets are converted from *RGB* to *HSI* and *YCbCr* colour spaces, and the data set is divided into training and test sets. In this setup the training set contains n images per subject, and the rest of the images, are used for the test set.

TABLE 7: Comparison of the proposed SVD based equalization with standard Histogram Equalization (HE) on the final recognition rates, where there are 5 poses in the training set.

Equalization methods	FERET Performance (%)		HP Performance (%)	
	MV	FVF	MV	FVF
SVD Based	94.20	98.00	97.33	97.78
HE	18.96	24.00	41.07	44.00

5. Fusion of Decision in Different Colour Channels

The face recognition procedure explained in the previous section can be applied to different colour channels such as H, S, I, Y, Cb, and Cr. Hence, given a face image the image can be represented in these colour spaces with dedicated colour PDFs for each channel. Different colour channels contain different information regarding the image; therefore all of these six PDFs can be combined to represent a face image. There are many techniques to combine the resultant decision. In this paper, sum rule, median rule, max rule, product rule, majority voting, and feature vector fusion methods have been used to do this combination [25].

These data fusion techniques use probability of the decisions they provide through classifiers. That is why it is necessary to calculate the probability of the decision of each classifier based on the minimum KLD value. This is achieved by calculating the probability of the decision in each colour channel, κ_C, which can be formulated as follows:

$$\varsigma_C = \frac{[\kappa_1 \ \kappa_2 \ \cdots \ \kappa_{nM}]_C}{\sum_{i=1}^{nM} \kappa_i}, \qquad K_C = \max(1 - \varsigma_C)$$
$$C = \{H, S, I, Y, Cb, Cr\} \tag{18}$$

where ς_C is the normalized KLD values, n shows the number of face samples in each class, and M is the number of classes. The highest similarity between two projection vectors is when the minimum KLD value is zero. This represents a prefect match, that is, the probability of selection is 1. So zero Euclidean distance represents probability of 1 that is why ς_C has been subtracted from 1. The maximum probability corresponds to the probability of the selected class. The sum rule is applied, by adding all the probabilities of a class in different colour channels followed by declaring the class with the highest accumulated probability to be the selected class. The maximum rule, as its name implies, simply takes the maximum among the probabilities of a class in different colour channels followed by declaring the class with the highest probability to be the selected class. The median rule is similarly takes the median among the sorted probabilities of a class in different channels. The product rule is achieved from the product of all probabilities of a class in different colour channels. It is very sensitive as a low probability (close to 0) will remove any chance of that class being selected [25].

Majority voting (MV) is another data fusion technique. The main idea behind MV is to achieve increased recognition rate by combining decisions of different colour channels. The MV procedure can be explained as follows. Given the probability of the decisions, κ_C, in all colour channels (C :

H, S, I, Y, Cb, Cr), the highest repeated decision among all channels is declared to be the overall decision.

Data fusion is not the only way to improve the decision making. PDFs vectors can also be simply concatenated with the feature vector fusion (FVF) process which is a source fusion technique and can be explained as follows. Consider $\{p_1, p_2, \ldots, p_M\}_C$ to be a set of training face images in colour channel C (H, S, I, Y, Cb, Cr), then for a given query face image, the fvf_q is defined as a vector which is the combination of all PDFs of the query image q as follow:

$$fvf_q = \begin{bmatrix} q_H & q_S & q_I & q_Y & q_{Cb} & q_{Cr} \end{bmatrix}_{1 \times 1536}. \tag{19}$$

This new PDF can be used to calculate the KLD between fvf_q and fvf_{p_i} of the images in the training samples as follows:

$$\chi_r = \min\left(\kappa\left(fvf_q, fvf_{p_i}\right)\right), \quad i = 1, \ldots, M \tag{20}$$

where M is the number of images in the training set. Thus, the similarity of the rth images in the training set and the query face can be reflected by χ_r, which is the minimum KLD value. The image is with the lowest KLD distance in a channel; χ_r is declared to be the vector representing the recognized subject. With the proposed system using PDFs in different colour channels as the face feature vector, discussed ensemble-based systems in decision making have been tested on the FERET, the Essex University, the Georgia tech university, and the HP face databases. The correct recognition rates in percent are included in Table 4. Each result is the average of 100 runs, where we have randomly shuffled the faces in each class.

6. Results and Discussions

The experimental results have been achieved by testing the system on the following face databases: The HP face database containing 150 faces of 15 classes with 10 different rotational poses varying from $-90°$ to $+90°$ for each class, a subset of the FERET face database containing 500 faces of 50 classes with 10 different poses varying from $-90°$ to $+90°$ for each class, the Essex University face database containing 1500 faces of 150 classes with 10 different slightly varying poses and illumination changes, and the Georgia Tech University face database containing 500 faces of 50 classes with 10 different varying poses, illumination, and background. The correct recognition rates in percent of the aforementioned face databases using PDF-based face recognition system in different colour channels are shown in Table 2. Each result is the average of 100 runs, where we have randomly shuffled

the faces in each class. It is important to note that the performance of each colour channel is different, which means that a person can be recognized in one channel where the same person may fail to be recognized in another channel.

In order to show the superiority of proposed PDF-based face recognition over PCA-based face recognition in each colour channel, the performance of PCA-based face recognition system on the aforementioned face databases in different colour channels is shown in Table 3.

The results of the proposed system using data and source fusion techniques, for different face databases have been shown in Table 4. The results show that the performance of the product rule dramatically drops when the number of images per subject in the training set is increasing, this is because by increasing the number of training images per subject, the probability of having a low probability will be increased, so one low probability is enough to cancel the effect of several high probabilities. The median rule is marginally better than sum rule in some occasion but from computational complexity point of view the median rule is more expensive than the sum rule, because it requires sorting. The marginal improvement of the median rule is due to this fact that having only one out of range probability will not affect the median, though it will affect the sum rule. The minimum rule has not been discussed in the work, as it is not logical to give priority to the decisions which have a low probability of occurrence. The same data fusion techniques have been applied to the PCA-based system in different colour channels to improve the final recognition rate. The recognition rates have been stated in Table 5. A comparison between Table 4 and Table 5 indicates the high performance of the proposed system.

In order to show the superiority of the proposed method on available state-of-art and conventional face recognition systems, we have compared the recognition rate with conventional PCA-based face recognition system and state-of-art techniques such as Nonnegative Matrix Factorization (NMF) [26, 27], supervised incremental NMF (INMF) [28], LBP [8], and LDA [3] based face recognition systems for the FERET face database. The experimental results are shown in Table 6. In Figure 4, the graphical illustration of the superiority of the proposed data fusion boosted colour PDF-based face recognition system over the aforementioned face recognition systems. Performance was achieved on FERET face database by two selected data fusion techniques FVF and median rule. The results clearly indicate that this superiority is achieved by using PDF-based face recognition in different colour channels backed by the data fusion techniques.

In an attempt to show the effectiveness of the proposed SVD-based equalization technique, the comparison between the proposed method and HE on the final recognition scores is shown in Table 7. As the results indicate, HE is not a suitable preprocessing technique for the proposed face recognition system, due to the fact that it transforms the input image such that the PDF of the output image has uniform distribution. This process dramatically reshapes the PDFs of the segmented face images, which results in poor recognition performance.

7. Conclusion

In this paper we introduced a high performance face recognition system based on combining the decision obtained from PDFs in different colour channels. A new preprocessing procedure was employed to equalize the images. Furthermore local SMQT technique has been employed to isolate the faces from the background, and KLD-based PDF matching is used to perform face recognition. Minimum KLD between the PDF of a given face and the PDFs of the faces in the database was used to perform the PDF matching. Several decision making techniques such as sum rule, minimum rule, median rule and product rule, majority voting, and feature vector fusion have been employed to improve the performance of the proposed PDF-based system. The performance clearly shows the superiority of the proposed system over the conventional and the state-of-art based face recognition systems.

References

[1] W. W. Bledsoe, "The model method in facial recognition," Tech. Rep. PRI 15, Panoramic Research, Palo Alto, Calif, USA, 1964.

[2] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 586–591, Maui, Hawaii, USA, June 1991.

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[4] S. Marcel and S. Bengio, "Improving face verification using skin color information," in *Proceedings of the 16th International Conference on Pattern Recognition*, vol. 2, pp. 378–381, August 2002.

[5] I. Laptev, "Improvements of object detection using boosted histograms," in *British Machine Vision Conference (BMVC '06)*, vol. 3, pp. 949–958, 2006.

[6] T.-W. Yoo and I.-S. Oh, "A fast algorithm for tracking human faces based on chromatic histograms," *Pattern Recognition Letters*, vol. 20, no. 10, pp. 967–978, 1999.

[7] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *Proceedings of the European Conference on Computer Vision*, vol. 3021 of *Lecture Notes in Computer Science*, pp. 469–481, 2004.

[8] Y. Rodriguez and S. Marcel, "Face authentication using adapted local binary pattern histograms," in *Proceedings of the 9th European Conference on Computer Vision (ECCV '06)*, vol. 3954 of *Lecture Notes in Computer Science*, pp. 321–332, Graz, Austria, May 2006.

[9] H. Demirel and G. Anbarjafari, "High performance pose invariant face recognition," in *Proceedings of the 3rd International Conference on Computer Vision Theory and Applications (VISAPP '08)*, vol. 2, pp. 282–285, Funchal, Portugal, January 2008.

[10] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34–58, 2002.

[11] D. Chai and K. N. Ngan, "Face segmentation using skin-color map in videophone applications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 4, pp. 551–564, 1999.

[12] M. Nilsson, J. Nordberg, and I. Claesson, "Face detection using local SMQT features and split up snow classifier," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '07)*, vol. 2, pp. 589–592, Honolulu, Hawaii, USA, April 2007.

[13] N. Gourier, D. Hall, and J. L. Crowley, "Estimating face orientation from robust detection of salient facial features," in *Proceedings of the Pointing, International Workshop on Visual Observation of Deictic Gestures (ICPR '04)*, Cambridge, UK, 2004.

[14] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.

[15] "Face Recognition Data," University of Essex, UK, The Data Archive, http://cswww.essex.ac.uk/mv/allfaces/index.html.

[16] "Face Recognition Data," Georgia Tech University, The Data Archive, December 2007, http://wwwanefian.com/research/face_reco.htm.

[17] M. Abdullah-Al-Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae, "A dynamic histogram equalization for image contrast enhancement," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 593–600, 2007.

[18] H. Demirel and G. Anbarjafari, "Pose invariant face recognition using probability distribution functions in different color channels," *IEEE Signal Processing Letters*, vol. 15, pp. 537–540, 2008.

[19] H. Demirel, G. Anbarjafari, and M. N. S. Jahromi, "Image equalization based on singular value decomposition," in *Proceedings of the 23rd International Symposium on Computer and Information Sciences (ISCIS '08)*, Istanbul, Turkey, 2008.

[20] Y. Tian, T. Tan, Y. Wang, and Y. Fang, "Do singular values contain adequate information for face recognition?" *Pattern Recognition*, vol. 36, no. 3, pp. 649–655, 2003.

[21] M. Nilsson, "Face detector software," provided in MathWorks exchange file, January 2008, http://www.mathworks.com/matlabcentral/fileexchange.

[22] M. Nilsson, M. Dahl, and I. Claesson, "The successive mean quantization transform," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 4, pp. 429–432, Philadelphia, Pa, USA, March 2005.

[23] E. Marszalec, B. Martinkauppi, M. Soriano, and M. Pietikäinen, "A physics-based face database for color research," *Journal of Electronic Imaging*, vol. 9, no. 1, pp. 32–38, 2000.

[24] S. Stanczak and H. Boche, "Information theoretic approach to the perron root of nonnegative irreducible matrices," in *Proceedings of IEEE Information Theory Workshop (ITW '04)*, pp. 254–259, October 2004.

[25] R. Polikar, "Ensemble based systems in decision making," *IEEE Circuits and Systems Magazine*, vol. 6, no. 3, pp. 21–45, 2006.

[26] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[27] D. D. Lee and H. S. Seung, "Algorithm for nonnegative matrix factorization," in *Proceedings of the Advances in Natural Information Processing System (NIPS '01)*, vol. 13, pp. 556–562, 2001.

[28] W.-S. Chen, B. Pan, B. Fang, M. Li, and J. Tang, "Incremental nonnegative matrix factorization for face recognition," *Mathematical Problems in Engineering*, vol. 2008, Article ID 410674, 17 pages, 2008.