



# Color space normalization: Enhancing the discriminating power of color spaces for face recognition

Jian Yang<sup>a,\*</sup>, Chengjun Liu<sup>b</sup>, Lei Zhang<sup>c</sup>

<sup>a</sup> School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, PR China

<sup>b</sup> Department of Computer Science, New Jersey Institute of Technology, Newark, NJ 07102, USA

<sup>c</sup> Department of Computing, Hong Kong Polytechnic University, Kowloon, Hong Kong

## ARTICLE INFO

### Article history:

Received 22 July 2008

Received in revised form

28 October 2009

Accepted 14 November 2009

### Keywords:

Color space

Color model

Face recognition

Face recognition grand challenge (FRGC)

Fisher linear discriminant analysis (FLD or LDA)

Biometrics

Pattern recognition

## ABSTRACT

This paper presents the concept of color space normalization (CSN) and two CSN techniques, i.e., the within-color-component normalization technique (CSN-I) and the across-color-component normalization technique (CSN-II), for enhancing the discriminating power of color spaces for face recognition. Different color spaces usually display different discriminating power, and our experiments on a large scale face recognition grand challenge (FRGC) problem reveal that the RGB and XYZ color spaces are weaker than the  $I_1I_2I_3$ , YUV, YIQ, and LSLM color spaces for face recognition. We therefore apply our CSN techniques to normalize the weak color spaces, such as the RGB and the XYZ color spaces, the three hybrid color spaces XGB, YRB and ZRG, and 10 randomly generated color spaces. Experiments using the most challenging FRGC version 2 Experiment 4 with 12,776 training images, 16,028 controlled target images, and 8,014 uncontrolled query images, show that the proposed CSN techniques can significantly and consistently improve the discriminating power of the weak color spaces. Specifically, the normalized RGB, XYZ, XGB, and ZRG color spaces are more effective than or as effective as the  $I_1I_2I_3$ , YUV, YIQ and LSLM color spaces for face recognition. The additional experiments using the AR database validate the generalization of the proposed CSN techniques. We finally explain why the CSN techniques can improve the recognition performance of color spaces from the color component correlation point of view.

© 2009 Elsevier Ltd. All rights reserved.

## 1. Introduction

Color provides an important clue or useful feature for object detection, tracking and recognition, image (or video) segmentation, indexing and retrieval, etc. [1–15]. Different color spaces (or color models) possess different characteristics and are suitable for different visual tasks. For instance, the *HSV* color space and the  $YC_bC_r$  color space are effective for face detection [2,3], while the modified  $L^*u^*v^*$  color space is useful for image segmentation [7]. As a result, when applying color information, we should first choose an appropriate color space, and such a choice is very important for achieving the best result for a specific visual task [15].

The RGB color space is a fundamental and widely used color space, and other color spaces (or color models) are usually defined by transformations of the RGB color space. The transformations involved are either linear or nonlinear. The color spaces generated via the nonlinear transformations (of the RGB color space), such as

the HSV and  $L^*a^*b^*$  color spaces [16], generally associate with the human vision system, while the color spaces determined by the linear transformations, such as the YUV and YIQ color spaces [17], usually associate with color display of some hardware (such as television and color monitors) for adapting to human color-response characteristics.

Although color has been demonstrated helpful for face detection and tracking, some past research suggests that color appears to confer no significant face recognition advantage beyond the luminance information [18]. Recent research efforts, however, reveal that color may provide useful information for face recognition. The experimental results in [19] show that the principal component analysis (PCA) method using color information can improve the recognition rate compared to the same method using only luminance information. The results in [20] reveal that color cues do play a role in face recognition and their contribution becomes evident when shape cues are degraded. The results in [45] further demonstrate that color cues can significantly improve recognition performance compared with intensity-based features for coping with low-resolution face images. Other research findings also demonstrate the effectiveness of color for face recognition [21–28].

\* Corresponding author.

E-mail addresses: csjyang@mail.njust.edu.cn (J. Yang), chengjun.liu@njit.edu (C. Liu), cszhang@comp.polyu.edu.hk (L. Zhang).

Different color spaces derived from different transformations of the RGB color space revealed different face recognition performance. The YUV color space, for example, is shown more effective than the RGB color space [19]. The  $YQ_C_r$  color configuration (a hybrid color space), where the Y and Q color components are from the YIQ color space and the  $C_r$  color component is from the  $Y C_b C_r$  color space, is more powerful than the RGB, HSV and  $L^*a^*b^*$  color spaces [23–25]. Another two hybrid color spaces,  $RIQ$  [44],  $RQ_C_r$  [45] are demonstrated effective recently. Some color spaces generated by evolution algorithms [43] and discriminant models [27,28] also turn out to be very powerful. Current research findings showed that some linear color spaces, which are derived by linear transformations from the RGB color space, perform much better those derived by nonlinear transformations from the RGB color space. We therefore focus on linear color spaces in this paper. Rather than searching for a more effective color space as the previous research [23–28,43–45], we try to explore general ways for enhancing the performance of conventional color spaces for face recognition.

This paper assesses the performance of different color spaces using a large scale database, the Face Recognition Grand Challenge (FRGC) version 2 database [29–31]. The assessment results reveal that some color spaces, such as the RGB, XYZ, HSV and  $L^*a^*b^*$  color spaces [16], are relatively weak, whereas the other color spaces, such as the  $I_1I_2I_3$  [32], YUV, YIQ and LSLM color spaces [33], are relatively powerful in achieving good face recognition performance. What characteristics make the  $I_1I_2I_3$ , YUV, YIQ and LSLM color spaces more powerful than the RGB and XYZ color spaces for face recognition? By analyzing the transformation matrices of the  $I_1I_2I_3$ , YUV, YIQ and LSLM colors spaces, we find out that these matrices all share a common characteristic: the sums of the elements in the second and third rows of the transformation matrix are both zero. The RGB and XYZ color spaces, however, do not have such a property. Inspired by the finding of the difference of the transformation matrices between the weak and powerful color spaces, we present the concept of color space normalization (CSN) and develop two CSN techniques. These CSN techniques normalize any color space that is derived by a linear transformation of the RGB color space, so that the normalized color space possesses the same properties as the powerful color spaces do, i.e., the sums of the elements in the second and third rows of the transformation matrix are both zero. The proposed two CSN techniques are demonstrated to be very effective: the normalized RGB and XYZ color spaces are as powerful as or even more powerful than the  $I_1I_2I_3$ , YUV, YIQ and LSLM color spaces for face recognition.

The proposed CSN techniques, which are capable of converting weak color spaces into powerful ones, provide us more flexibility for color space selection for specific pattern recognition tasks. Previous color space selection is limited to the set of conventional color spaces or their hybrids. Specifically, we choose a powerful color space by experiments from two sets of color spaces: one set of conventional color spaces and one set of hybrid color spaces that are generated by choosing some color components from the conventional color spaces. The weak color spaces are simply left behind due to their unsatisfactory performance. The proposed color space normalization techniques, however, can convert the weak color spaces into powerful ones, and these normalized color spaces form a new set of color spaces, from which we might find a more effective color space for a specific recognition task. The three sets of color spaces are illustrated in Fig. 1.

The remainder of paper is organized as follows. Section 2 outlines some conventional color spaces. Section 3 presents the concept of color space normalization (CSN) and two CSN techniques. In Section 4, the proposed CSN techniques are assessed using the face recognition grand challenge (FRGC)

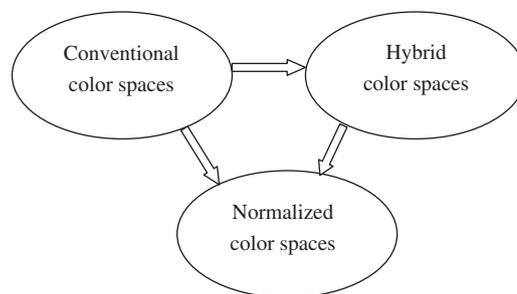


Fig. 1. Illustration of three sets of color spaces.

database as well as the AR database, and the problem of why the proposed CSN techniques can improve the face verification and recognition performance is addressed. Finally, some conclusions are offered in Section 5.

## 2. Conventional color spaces

A color space, generally associated with the definition of a color model, is a means of uniquely specifying a color. There are a number of color spaces in common usage depending on the particular application involved. The RGB color model is an additive model in which red, green, and blue are combined in various ways to reproduce different colors. The RGB color space, therefore, describes color by three components: red, green, and blue. The RGB color space is a fundamental and commonly used color space. Other color spaces can be calculated from the RGB color space by means of either linear or nonlinear transformations. It is apparent that every color space derived by the linear transformation of the RGB color space is uniquely determined by the associated transformation matrix. In the following, we review five color spaces derived from the RGB color space via linear transformations.

The XYZ color space was derived from a series of experiments in the study of the human perception by the International Commission on Illumination (CIE) in 1931. The transformation from the RGB color space to the XYZ color space is as follows [16]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.607 & 0.174 & 0.201 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.117 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

The  $I_1I_2I_3$  color space was obtained through the decorrelation of the RGB color components using the  $K$ - $L$  transform by Ohta et al. [32] in 1980. The transformation from the RGB color space to the  $I_1I_2I_3$  color space is as follows:

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 0 & -1/2 \\ -1/2 & 1 & -1/2 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2)$$

The YUV color space is defined in terms of one luminance (Y) and two chrominance components (U and V), and it is used in the PAL (phase alternating line), NTSC (National Television System Committee), and SECAM (Séquentiel couleur à mémoire) composite color video standards [17]. The transformation from the RGB color space to the YUV color space is as follows:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1471 & -0.2888 & 0.4359 \\ 0.6148 & -0.5148 & -0.1000 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3)$$

The YIQ color space was formerly used in the National Television System Committee (NTSC) television standard [17]. The YIQ system, which is intended to take advantage of human

color-response characteristics, and can be derived from the corresponding RGB space as follows:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ 0.5957 & -0.2744 & -0.3213 \\ 0.2115 & -0.5226 & 0.3111 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4)$$

Note that the  $I$  and  $Q$  components in the YIQ color space are obtained via a clockwise rotation ( $33^\circ$ ) of the  $U$  and  $V$  color components in the YUV color space.

The LSLM color space is a linear transformation of the RGB color space based on the opponent signals of the cones: black–white, red–green, and yellow–blue. The LSLM color space is defined as follows [33]:

$$\begin{bmatrix} L \\ S \\ LM \end{bmatrix} = \begin{bmatrix} 0.209 & 0.715 & 0.076 \\ 0.209 & 0.715 & -0.924 \\ 3.148 & -2.799 & -0.349 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5)$$

For a given color image, the color component images corresponding to the six color spaces discussed in this section are illustrated in Fig. 2.

### 3. Color space normalization: concept and techniques

Different color spaces usually display different discriminating power, and our experiments on a large scale face recognition grand challenge (FRGC) problem reveal that some color spaces, such as the RGB and XYZ color spaces, are relatively weak, whereas other color spaces, such as the  $I_1I_2I_3$ , YUV, YIQ and LSLM color spaces, are relatively powerful (for more details, please refer to the experimental assessment in Section 4). What characteristics make the  $I_1I_2I_3$  [32], YUV, YIQ and LSLM color spaces more powerful than the RGB and XYZ color spaces for recognition? By analyzing the transformation matrices of the  $I_1I_2I_3$ , YUV, YIQ and LSLM colors spaces, we find out that these matrices all share a common characteristic: the sums of the elements in the second and third rows of the transformation matrix are both zero. The RGB and XYZ color spaces, however, do not have such a property.

Note that the transformation matrix of the RGB color space is an identity matrix:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (6)$$

Inspired by the finding of the difference of the transformation matrices between the weak and powerful color spaces, we present the concept of color space normalization (CSN) and develop two CSN techniques. These CSN techniques normalize any color space that is derived by a linear transformation of the RGB color space, so that the normalized color space possesses the same property as the powerful color spaces do, i.e., the sums of the elements in the second and third rows of the transformation matrix are both zero.

#### 3.1. Within-color-component normalization

To achieve the goal that the sums of the elements in the second and the third rows of the color space transformation matrix are zero, the within-color-component normalization technique works by directly removing the means of the second and the third row vectors, respectively. Let  $C_1$ ,  $C_2$  and  $C_3$  be the three color components derived by the following linear transformation of the RGB color space:

$$\begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix} = \mathbf{A} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \mathbf{A}_3 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (7)$$

The mean of the second row vector of the transformation matrix  $\mathbf{A}$  is  $m_2 = (a_{21} + a_{22} + a_{23})/3$  and the mean of the third row vector is  $m_3 = (a_{31} + a_{32} + a_{33})/3$ . Removing  $m_2$  from the second row vector and  $m_3$  from the third row vector, we obtain a normalized transformation matrix  $\bar{\mathbf{A}}_1$ , which determine the normalized color

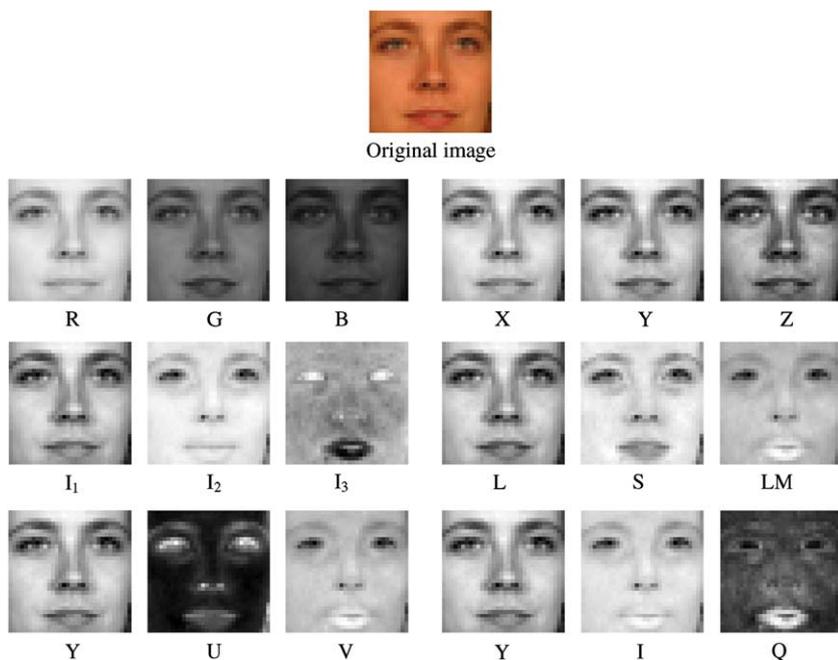


Fig. 2. Illustration of the three color component images corresponding to the six color spaces.

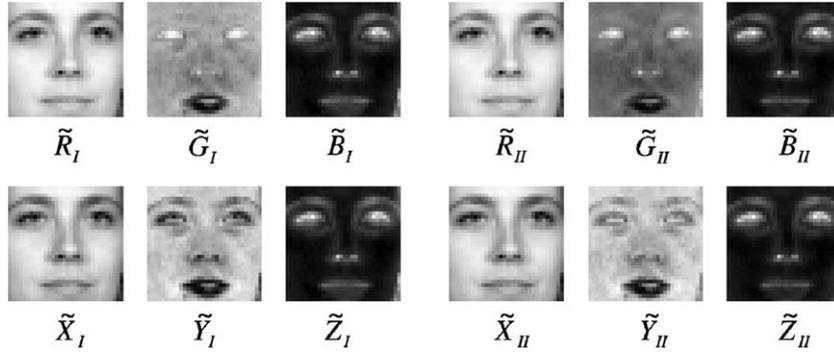


Fig. 3. Illustration of the color component images of the normalized RGB and XYZ color spaces using CSN-I and CSN-II, respectively.

space  $\tilde{C}_1\tilde{C}_2\tilde{C}_3$ :

$$\begin{bmatrix} \tilde{C}_1 \\ \tilde{C}_2 \\ \tilde{C}_3 \end{bmatrix} = \tilde{\mathbf{A}}_I \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{A}}_1 \\ \tilde{\mathbf{A}}_2 \\ \tilde{\mathbf{A}}_3 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21}-m_2 & a_{22}-m_2 & a_{23}-m_2 \\ a_{31}-m_3 & a_{32}-m_3 & a_{33}-m_3 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (8)$$

The within-color-component normalization technique is named color space normalization I (CSN-I).

For example, the normalized RGB color space using CSN-I is

$$\begin{bmatrix} \tilde{R}_I \\ \tilde{G}_I \\ \tilde{B}_I \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (9)$$

The normalized XYZ color space using CSN-I is

$$\begin{bmatrix} \tilde{X}_I \\ \tilde{Y}_I \\ \tilde{Z}_I \end{bmatrix} = \begin{bmatrix} 0.6070 & 0.1740 & 0.2000 \\ -0.0343 & 0.2537 & -0.2193 \\ -0.3940 & -0.3280 & 0.7220 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (10)$$

For the original color image shown in Fig. 2, the color component images corresponding to these two normalized color spaces are illustrated in Fig. 3.

### 3.2. Across-color-component normalization

To make the sums of the elements in the second and the third rows of the color space transformation matrix be zero, the across-color-component normalization technique works in the following way. The original three row vectors of the color space transformation matrix are first used to generate two zero-mean row vectors via a linear combination. A new color space transformation matrix is then obtained by replacing the second and third row vectors of the original transformation matrix with the generated two zero-mean row vectors.

The linear combination of the three row vectors of the original color space transformation matrix  $\mathbf{A}$  may be written as follows:

$$\xi = k_1\mathbf{A}_1 + k_2\mathbf{A}_2 + k_3\mathbf{A}_3 = \left( \sum_{i=1}^3 k_i a_{i1}, \sum_{i=1}^3 k_i a_{i2}, \sum_{i=1}^3 k_i a_{i3} \right) \quad (11)$$

Let the sum of the elements of this linear combination vector  $\xi$  (row vector) be zero, i.e.,

$$\begin{aligned} \sum_{i=1}^3 k_i a_{i1} + \sum_{i=1}^3 k_i a_{i2} + \sum_{i=1}^3 k_i a_{i3} &= k_1 \sum_{j=1}^3 a_{1j} + k_2 \sum_{j=1}^3 a_{2j} + k_3 \sum_{j=1}^3 a_{3j} \\ &= [s_1, s_2, s_3][k_1, k_2, k_3]^T = 0, \end{aligned} \quad (12)$$

where  $s_i = \sum_{j=1}^3 a_{ij}$ ,  $i = 1, 2, 3$ . Obviously,  $s_i$  is the sum of the elements of the  $i$ -th row vector of the color space transformation matrix  $\mathbf{A}$ .

Eq. (12) shows that the linear combination coefficient vector  $[k_1, k_2, k_3]^T$  can be chosen as the basis vectors of the null space of  $[s_1, s_2, s_3]$ . Since this null space is two-dimensional, it has only two basis vectors. Let the two basis vectors<sup>1</sup> be  $\mathbf{K}_1 = [k_{11}, k_{21}, k_{31}]^T$  and  $\mathbf{K}_2 = [k_{12}, k_{22}, k_{32}]^T$ . We then obtain the two zero-mean row vectors as follows:

$$\xi_1 = \mathbf{K}_1^T \mathbf{A} = k_{11}\mathbf{A}_1 + k_{21}\mathbf{A}_2 + k_{31}\mathbf{A}_3 = \left( \sum_{i=1}^3 k_{i1} a_{i1}, \sum_{i=1}^3 k_{i1} a_{i2}, \sum_{i=1}^3 k_{i1} a_{i3} \right) \quad (13)$$

$$\xi_2 = \mathbf{K}_2^T \mathbf{A} = k_{12}\mathbf{A}_1 + k_{22}\mathbf{A}_2 + k_{32}\mathbf{A}_3 = \left( \sum_{i=1}^3 k_{i2} a_{i1}, \sum_{i=1}^3 k_{i2} a_{i2}, \sum_{i=1}^3 k_{i2} a_{i3} \right) \quad (14)$$

The normalized color space transformation matrix is defined as follows:

$$\tilde{\mathbf{A}}_{II} = \begin{bmatrix} \mathbf{A}_1 \\ \xi_1 \\ \xi_2 \end{bmatrix}, \quad (15)$$

which determines the following normalized color space  $\tilde{C}_1\tilde{C}_2\tilde{C}_3$ :

$$\begin{bmatrix} \tilde{C}_1 \\ \tilde{C}_2 \\ \tilde{C}_3 \end{bmatrix} = \tilde{\mathbf{A}}_{II} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (16)$$

The across-color-component normalization technique is named color space normalization II (CSN-II).

For example, the normalized RGB color space using CSN-II is

$$\begin{bmatrix} \tilde{R}_{II} \\ \tilde{G}_{II} \\ \tilde{B}_{II} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -0.5774 & 0.7887 & -0.2113 \\ -0.5774 & -0.2113 & 0.7887 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (17)$$

The normalized XYZ color space using CSN-II is

$$\begin{bmatrix} \tilde{X}_{II} \\ \tilde{Y}_{II} \\ \tilde{Z}_{II} \end{bmatrix} = \begin{bmatrix} 0.6070 & 0.1740 & 0.2000 \\ -0.0901 & 0.3631 & -0.2730 \\ -0.4600 & -0.1986 & 0.6586 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (18)$$

For the original color image shown in Fig. 2, the color component images corresponding to these two normalized color spaces are illustrated in Fig. 3.

<sup>1</sup> The two basis vectors are not necessarily orthogonal. But, here we prefer the orthogonal basis vectors and use the matlab function `null([s1, s2, s3])` to calculate them.

### 3.3. An insight into the normalized color spaces

To facilitate better understanding of the proposed color space normalization techniques (especially for CSN-II: the across-color-component normalization), we provide more details of some color spaces (models) defined by NTSC. In the YUV color space, for example, the luminance component,  $Y$ , is derived from the XYZ color space, while the remaining two chrominance components,  $U$  and  $V$ , are derived from the linear combination of the  $Y$  component, and the  $R$  and  $B$  components from the RGB color space. Specifically, let us first define the YRB color space as follows:

$$\begin{bmatrix} Y \\ R \\ B \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (19)$$

From the YRB color space, the  $C$ - $Y$  color model (space) is defined as follows [16]:

$$\begin{bmatrix} Y \\ R-Y \\ B-Y \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.701 & -0.587 & -0.114 \\ -0.299 & -0.587 & 0.866 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (20)$$

The YUV color space is a scaled version of the  $C$ - $Y$  color space, and the  $U$  and  $V$  components are computed as follows:

$$U = \frac{0.436}{0.866}(B-Y) \quad \text{and} \quad V = \frac{0.615}{0.701}(R-Y) \quad (21)$$

The YUV color space is equivalent to the  $C$ - $Y$  color space for face recognition because the scaling factors have been eliminated (by normalizing the standard deviation of each component image) in our experiments in Section 4.

The  $C$ - $Y$  color space, as a matter of fact, can be viewed as a normalized YRB color space using the across-color-component normalization technique. For the transformation matrix  $\mathbf{A}$  of the YRB color space, the sum of the elements in each row is one, i.e.,  $[s_1, s_2, s_3] = [1, 1, 1]$ . As a result, the linear equation in Eq. (12) has two basis vectors:  $\mathbf{K}_1 = [-1, 1, 0]^T$  and  $\mathbf{K}_2 = [-1, 0, 1]^T$ . These two basis vectors determine the two zero-mean row vectors,  $\xi_1 = \mathbf{K}_1^T \mathbf{A} = R - Y$  and  $\xi_2 = \mathbf{K}_2^T \mathbf{A} = B - Y$ , respectively. Thus, the transformation matrix of the normalized YRB color space is

$$\begin{bmatrix} Y \\ R-Y \\ B-Y \end{bmatrix}$$

which is exactly the transformation matrix of the  $C$ - $Y$  color space. Note that these two basis vectors  $\mathbf{K}_1 = [-1, 1, 0]^T$  and  $\mathbf{K}_2 = [-1, 0, 1]^T$  are not orthogonal.

Now, we can interpret the normalized color space  $\tilde{C}_1 \tilde{C}_2 \tilde{C}_3$  from the color model (space) point of view, that is,  $\tilde{C}_1$  can be viewed as the luminance component while  $\tilde{C}_2$  and  $\tilde{C}_3$  can be viewed as the chrominance components.

## 4. Experimental evaluation and analysis

This section first presents the face recognition grand challenge (FRGC) database, the evaluation criteria and methodology, and then assesses the performance of six conventional color spaces and a series of normalized color spaces. All experimental results show that the proposed CSN techniques improve the discriminating power of the color spaces. This section finally provides an explanation of why the CSN techniques can improve the performance of color spaces for recognition.

### 4.1. Database, evaluation criteria and methodology

#### 4.1.1. The FRGC and AR databases

We assess the performance of different color spaces using a large scale database, the FRGC version 2 database [29,30]. This database contains 12,776 training images, 16,028 controlled target images, and 8,014 uncontrolled query images for the FRGC Experiment 4. The controlled images have good image quality, while the uncontrolled images display poor image quality, such as large illumination variations, low resolution of the face region, and possible blurring. It is these uncontrolled factors that pose the grand challenge to face recognition performance [23]. The BEE system [21] provides a computational experimental environment to support a challenge problem in face recognition, and it allows the description and distribution of experiments in a common format. The BEE system uses the PCA method that has been optimized for large scale problems as a baseline algorithm, which applies the whitened cosine similarity measure [21]. The BEE baseline algorithm shows that the FRGC Experiment 4, which is designed for indoor controlled single still image versus uncontrolled single still image, is the most challenging FRGC experiment. We therefore choose the FRGC Experiment 4 to evaluate our method. In our experiments, the face region of each image is first cropped from the original high-resolution still images and resized to a spatial resolution of  $32 \times 32$ . Fig. 4 shows some example FRGC images used in our experiments. Following the FRGC protocol, we use the standard training set of the FRGC version 2 Experiment 4 for training.

We further validate the effectiveness and generalization of the proposed method using another database: the AR database [40]. The AR database contains over 4,000 color face images of 126 people (70 men and 56 women), including frontal views of faces with different facial expressions, lighting conditions and occlusions. The pictures of most persons were taken in two sessions (separated by two weeks). Each session contains 13 color images and 120 individuals (65 men and 55 women) participated in both sessions. The images of these 120 individuals were selected and used in our experiment. Only the full facial images were considered here (no attempt was made to handle occluded face recognition in each session). We manually cropped the face portion of the image and then normalized it to  $50 \times 45$  pixels. The normalized images of one person are shown in Fig. 5, where (a)–(g) are from Session 1, and (n)–(t) are from Session 2. The details of the images are: (a) neutral expression, (b) smile, (c) anger, (d) scream, (e) left light on; (f) right light on; (g) all sides light on; and (n)–(t) were taken under the same conditions as (a)–(g). In our experiments, images from the first session (i.e., (a)–(g)) were used for training, and images from the second session (i.e., (n)–(t)) were used for testing.

#### 4.1.2. Evaluation criteria

Two criteria are involved to evaluate the performance of different color spaces: one is the *verification rate* and the other is the *recognition rate*. The FRGC protocol recommends using the receiver operating characteristic (ROC) curves, which plot the face verification rate (FVR, i.e., the true accept rate) versus the false accept rate (FAR), to report the face recognition performance. The ROC curves are automatically generated by the BEE system when a similarity matrix is input to the system. In particular, the BEE system generates three ROC curves, ROC I, ROC II, and ROC III, corresponding to images collected within semesters, within a year, and between semesters, respectively. The face verification rate at the false accept rate of 0.1% is generally used as a standard for performance comparison.



Fig. 4. Example FRGC images that have been cropped to 32 × 32.

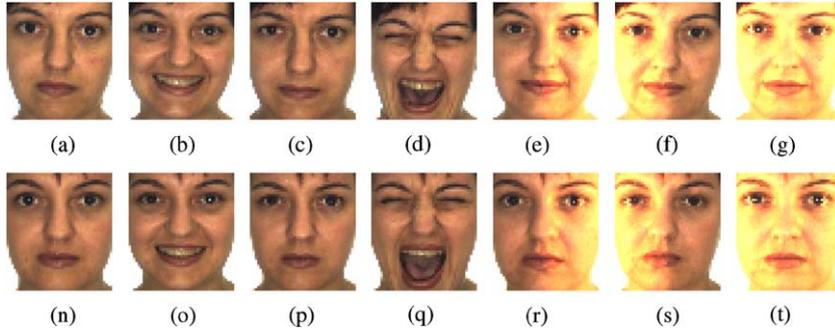


Fig. 5. Sample images for one subject of the AR database.

The recognition rate is another popular evaluation criterion for face recognition, although it is not recommended by the FRGC protocol. To obtain the recognition rate, we need to first calculate distance between every pair of a query image and a target image, and then use the nearest-neighbor classifier for classifying all query images. The recognition rate is the ratio of the number of correctly classified query images to the total number of query images.

It should be pointed out that the verification rate in the FRGC Experiment 4 emphasizes the similarity of samples that are relatively distant from one another within a class because it needs to measure all similarity between samples, whereas the recognition rate emphasizes the similarity of samples that are close to one another within a class since it applies a nearest-neighbor classifier. Therefore, these two criteria evaluate the performance of different color spaces for recognition from different viewpoints.

#### 4.1.3. Evaluation methodology

We evaluate the performance of different color spaces using the Fisher linear discriminant analysis (FLD) method [34–38]. In the FLD method, the between-class scatter matrix  $\mathbf{S}_b$  and the within-class scatter matrix  $\mathbf{S}_w$  are defined as follows [39]:

$$\mathbf{S}_b = \frac{1}{M} \sum_{i=1}^c l_i (\mathbf{m}_i - \mathbf{m}_0)(\mathbf{m}_i - \mathbf{m}_0)^T \quad (22)$$

$$\mathbf{S}_w = \frac{1}{M} \sum_{i=1}^c \frac{l_i}{l_i - 1} \sum_{j=1}^{l_i} (\mathbf{x}_{ij} - \mathbf{m}_i)(\mathbf{x}_{ij} - \mathbf{m}_i)^T \quad (23)$$

where  $\mathbf{x}_{ij}$  denotes the  $j$ -th training sample in class  $i$ ;  $M$  is the total number of training samples,  $l_i$  is the number of training samples in class  $i$ ;  $c$  is the number of classes;  $\mathbf{m}_i$  is the mean of the training samples in class  $i$ ;  $\mathbf{m}_0$  is the mean across all training samples. The

Fisher discriminant basis vectors  $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \dots, \boldsymbol{\phi}_d$  are selected as the generalized eigenvectors of  $\mathbf{S}_b$  and  $\mathbf{S}_w$  corresponding to the  $d$  ( $d \leq c - 1$ ) largest generalized eigenvalues, i.e.,  $\mathbf{S}_b \boldsymbol{\phi}_j = \lambda_j \mathbf{S}_w \boldsymbol{\phi}_j$ , where  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ . To avoid overfitting and to improve generalization performance of the FLD method, the principal component analysis (PCA) method [34–38] is generally applied for dimensionality reduction in advance of the FLD method. This combined method is called the PCA+FLD method.

To combine the information in the three discriminating color component images for recognition purpose, we concatenate the three color component images  $C_1, C_2$  and  $C_3$  into one pattern vector and then perform PCA+FLD on the concatenated pattern vector. To avoid the negative effect of magnitude dominance of one component image over the others, we apply a basic image normalization method by removing the mean and normalizing the standard deviation of each component image before the concatenation. In the PCA+FLD method, we choose 1000 principal components after PCA and derive 220 discriminant features after FLD. We apply the cosine similarity measure to calculate the similarity scores of the query and target image pairs. The cosine similarity measure between two vectors  $\mathbf{x}$  and  $\mathbf{y}$  is defined as follows:

$$\delta_{\cos}(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \quad (24)$$

where  $\|\cdot\|$  is the notation of the Euclidean norm.

In the FRGC experiments, the similarity matrix stores the similarity score of every query image versus target image pair. As a result, the size of the similarity matrix is  $T \times Q$ , where  $T$  is the number of target images (16,028 for FRGC version 2 Experiment 4) and  $Q$  is the number of query images (8,014 for FRGC version 2 Experiment 4). The similarity matrix is input to the BEE system and three ROC curves (the verification rate vs. the false accept

rate) are generated. In addition, for achieving the recognition rate, we use the cosine distance, i.e.,  $-\delta_{\cos}(\mathbf{x}, \mathbf{y})$ , in the nearest-neighbor classifier for experiments on both databases.

#### 4.2. Experiments and results on the FRGC database

We first assess the performance of eight conventional color spaces and show that the RGB and XYZ color spaces are weaker than the  $I_1I_2I_3$ , YUV, YIQ, and LSLM color spaces for face recognition. We then design three experiments to evaluate the effectiveness of the proposed color space normalization (CSN) techniques for improving the performance of weak color spaces. In the first experiment, we normalize the RGB and XYZ color spaces. In the second experiment, we normalize the hybrid color spaces generated by combining the color components from the RGB and XYZ color spaces. In the third experiment, we normalize the randomly generated color spaces defined by random transformations of the RGB color space.

In the following experiments, for a color space  $C_1C_2C_3$ ,  $C_1C_2C_3$ -NI denotes its normalized version using CSN-I, and  $C_1C_2C_3$ -NII denotes its normalized version using CSN-II. For example, RGB-NI denotes the normalized RGB color space using CSN-I, and RGB-NII denotes the normalized RGB color space using CSN-II.

##### 4.2.1. Experiments using the eight conventional color spaces

We use the FRGC version 2 database, the evaluation criteria (recognition rate and verification rate), and the evaluation methodology (the PCA+LDA method) introduced in Section 4.1 to assess the six linear color spaces: RGB, XYZ,  $I_1I_2I_3$ , YUV, YIQ, and LSLM and two nonlinear color spaces: HSV and  $L^*a^*b^*$ . Fig. 6 shows the three ROC curves corresponding to each of these six linear color spaces. The performances of the hybrid color spaces YQCr [23] and RQCr [45], and the BEE baseline algorithm on grayscale images (obtained by averaging the  $R$ ,  $G$  and  $B$  components) are also included for comparison. Table 1 presents the recognition rates and the verification rates at the false accept rate of 0.1%. The recognition rate of the same evaluation method on the grayscale images is also listed in Table 1. These results indicate that the  $I_1I_2I_3$ , YUV, YIQ and LSLM color spaces have similar performance, except that the YUV color space is a little worse than the other three color spaces. The performance of these four color spaces is much better than that of the RGB, XYZ, HSV and  $L^*a^*b^*$  color spaces. Even the YUV color space achieves a verification rate over 9% higher than that achieved by the XYZ color space based on the ROC III curves, and achieves a recognition rate nearly 5% higher than that by the RGB color space. From these results, we can conclude that the RGB and XYZ color spaces are weaker than the  $I_1I_2I_3$ , YUV, YIQ and LSLM color spaces for face recognition. In addition, the nonlinear transformation generated color spaces, HSV and  $L^*a^*b^*$ , does not show any advantage over the RGB and XYZ color spaces.

##### 4.2.2. Experiments using the normalized RGB and XYZ color spaces

We now normalize the RGB and XYZ color spaces using two CSN techniques and obtain four normalized color spaces: RGB-NI, RGB-NII, XYZ-NI and XYZ-NII. We then apply the FRGC Experiment version 2 database, the evaluation criteria, and the evaluation methodology introduced in Section 4.1 to evaluate the performance of the four normalized color spaces. The experimental results are shown in Fig. 7 and Table 2. In particular, Fig. 7 shows the three ROC curves corresponding to the four normalized color spaces. The ROC curves of the RGB color space, the XYZ color space and the BEE baseline algorithm on grayscale images are also included in Fig. 7 for comparison. Table 2 presents the recognition rate and the verification rates at the

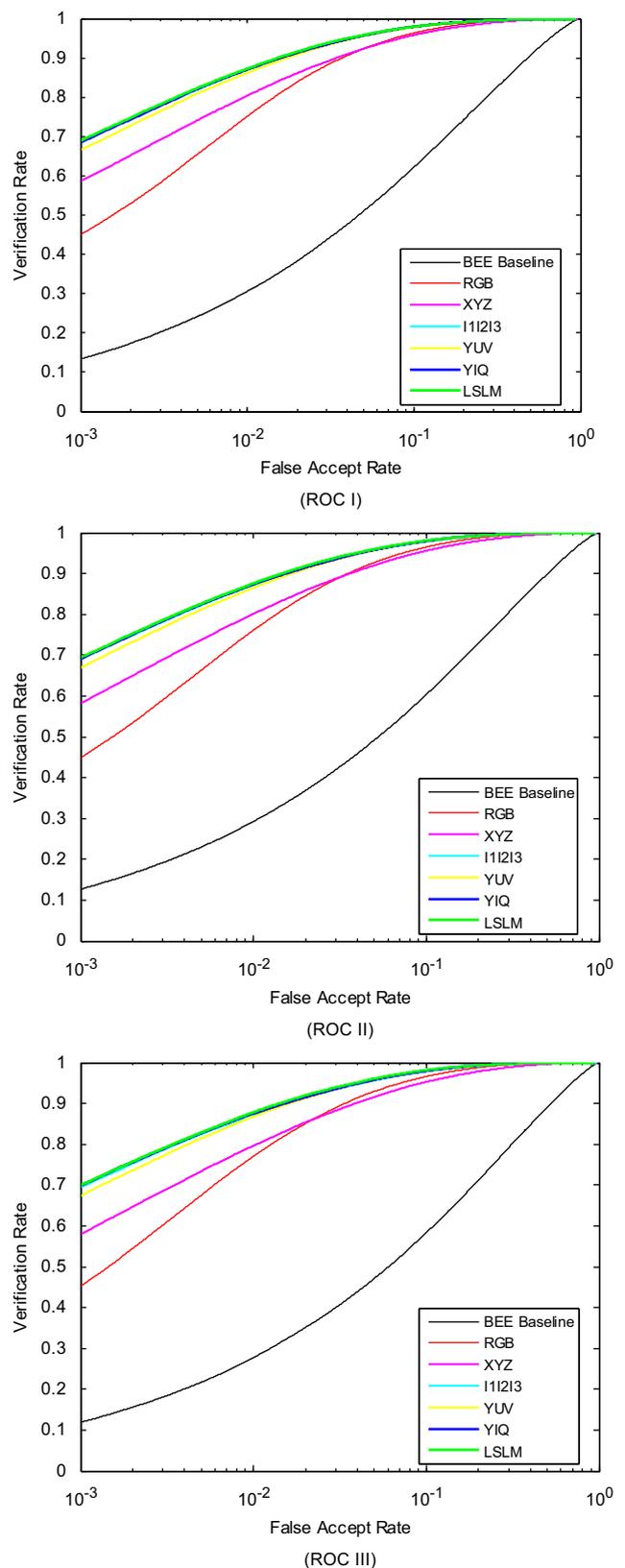


Fig. 6. The ROC curves corresponding to the six conventional color spaces and the BEE baseline performance. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

false accept rate of 0.1%. From Fig. 7 and Table 2, we can see that both the verification rates and the recognition rates are significantly improved after the color space normalization, no

**Table 1**  
Comparisons of the recognition rate (%) and the verification rate (%) at the false accept rate of 0.1%.

Color Space	Verification rate			Recognition rate
	ROC I	ROC II	ROC III	
RGB	44.95	44.80	45.20	89.1
XYZ	58.16	58.16	57.88	84.3
I <sub>1</sub> I <sub>2</sub> I <sub>3</sub>	68.58	68.95	69.36	94.5
YUV	66.40	66.81	67.29	94.0
YIQ	68.45	69.03	69.75	94.6
L <sub>SLM</sub>	68.84	69.36	69.88	94.3
YQCr (YQV)	71.12	71.54	71.85	94.7
RQCr	71.64	72.16	72.51	94.8
HSV	58.84	58.65	58.42	85.1
L <sup>*</sup> a <sup>*</sup> b <sup>*</sup>	45.21	45.25	45.59	86.5
Grayscale	36.88	36.65	36.94	76.2

matter which CSN technique is applied. Specifically, the results in Tables 1 and 2 show that the two normalized color spaces, RGB-NI and XYZ-NII, outperform those powerful conventional color spaces, the I<sub>1</sub>I<sub>2</sub>I<sub>3</sub>, YUV, YIQ and L<sub>SLM</sub> color spaces, in terms of the verification rate.

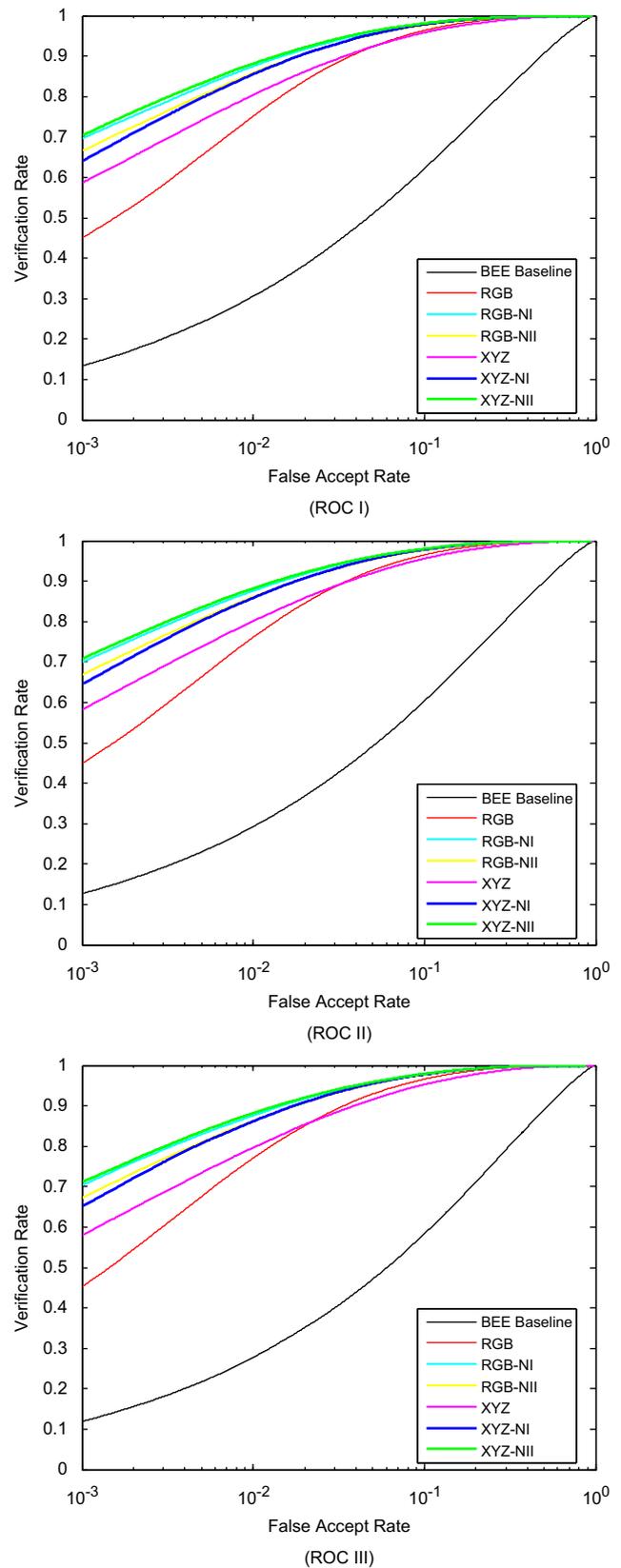
To gain more insights into color spaces, let us provide the performance of individual color components of the RGB and XYZ color spaces, as presented in Table 3. By comparing the results in Tables 2 and 3, we find that the combination of three color components is more effective than the use of one individual color component, especially for recognition rates. However, we can also find that a simple combination of three color components, for example, directly concatenating the R, G, and B components, cannot significantly improve the verification rate of the R component. This is due to the high correlation between the three color components. We will address this problem in detail in Section 4.3.

4.2.3. Experiments using Hybrid color spaces

Inspired by the formation of the YUV color space (which is derived from a hybrid color space YRB), we generate three hybrid color spaces, the XGB, YRB and ZRG color spaces, by configuring the components from the RGB and XYZ color spaces. Normalizing these three hybrid color spaces using the proposed CSN techniques can produce six normalized color spaces: XGB-NI, XGB-NII, YRB-NI, YRB-NII, ZRG-NI, and ZRG-NII. These color spaces are evaluated in the way as suggested in Section 4.1. The experimental results, as presented in Table 4, demonstrate again the effectiveness of the proposed normalization techniques. The recognition rate is improved by at least 5% and the verification rate is improved by more than 10% after the color space normalization. The ZRG-NII color space, for example, achieves the best verification rate (72.86% based on the ROC III curve) among all color spaces. Note that the performance of the YRB-NII color space is different from that of the YUV color space, although they are both derived from the hybrid YRB color space. This is because the generation of ZRG-NII is somewhat different from that of the YUV color space: an orthogonality constraint is imposed to calculate the basis vectors in the generation of the ZRG-NII color space.

4.2.4. Experiments using randomly generated color spaces

To further test the effectiveness of the proposed CSN techniques, we randomly generate 10 3 × 3 matrices using the matlab function “rand” (Note that the elements of each generated matrix are all between 0 and 1). Each of these matrices is used as a transformation matrix to determine a color space that is transformed from the RGB color space. The generated color space



**Fig. 7.** ROC curves corresponding to the RGB color space, the XYZ color space, and their corresponding normalized color spaces generated by the color space normalization techniques. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

is normalized using the proposed two CSN techniques. All color spaces and their normalized versions are evaluated in the way as suggested in Section 4.1. The corresponding verification rates and recognition rates are presented in Tables 5 and 6, respectively.

**Table 2**

Comparisons of the recognition rate (%) and the verification rate (%) at the false accept rate of 0.1% using the RGB, XYZ color spaces and their normalized color spaces.

Color space	Verification rate			Recognition rate
	ROC I	ROC II	ROC III	
RGB	44.95	44.80	45.20	89.1
RGB-NI	69.52	69.87	70.28	94.3
RGB-NII	66.41	66.76	67.07	94.4
XYZ	58.16	58.16	57.88	84.3
XYZ-NI	63.94	64.37	65.01	93.8
XYZ-NII	70.27	70.69	71.12	94.4

**Table 3**

Comparisons of the recognition rate (%) and the verification rate (%) at the false accept rate of 0.1% using the individual R, G, B, X, Y, Z color component images.

Color component image	Verification rate			Recognition rate
	ROC I	ROC II	ROC III	
R	43.54	43.92	44.58	82.8
G	35.89	35.56	35.48	74.3
B	27.83	27.12	26.24	61.3
X	47.49	47.79	48.17	80.1
Y	39.23	38.86	38.83	77.4
Z	27.82	27.06	26.21	62.0

**Table 4**

Comparisons of the recognition rate (%) and the verification rate (%) at the false accept rate of 0.1% using the hybrid color spaces and their normalized color spaces.

Color space	Verification rate			Recognition rate
	ROC I	ROC II	ROC III	
XGB	41.03	40.82	40.92	87.1
XGB-NI	70.54	70.60	70.62	94.2
XGB-NII	70.41	70.48	70.51	94.2
YRB	49.70	49.87	50.49	86.8
YRB-NI	61.08	61.67	62.37	92.9
YRB-NII	58.07	58.48	59.18	92.6
ZRG	42.72	42.82	43.19	89.0
ZRG-NI	68.03	68.22	68.29	94.0
ZRG-NII	73.00	72.94	72.86	94.4

**Table 5**

Verification rates (based on ROC III) of 10 randomly generated color spaces before color space normalization and after color space normalization.

No. of generated color space	Before CSN (ROC III)	After CSN-I (ROC III)	After CSN-II (ROC III)
1	52.32	73.52	70.28
2	47.87	58.96	56.34
3	35.31	72.13	56.33
4	45.62	63.17	57.36
5	45.67	68.96	67.70
6	49.27	63.04	71.01
7	53.50	64.14	63.83
8	48.91	69.46	69.07
9	53.16	65.37	65.16
10	51.18	68.76	70.71
Average	<b>48.28</b>	<b>66.75</b>	<b>64.78</b>

Note that in Table 5, we only list the verification rates (at the false accept rate of 0.1%) corresponding to ROC III for simplicity, since all the three ROCs have very similar performance. The results in Tables 5 and 6 show that the two CSN techniques (CSN-I and CSN-II) consistently improve the verification and recognition performance of the randomly generated color spaces by large margins. In particular, the average verification rate is improved 16% and the average recognition rate is improved 12%. These results demonstrate again the effectiveness of the proposed two CSN techniques.

#### 4.3. Experiments and results on the AR database

To validate the generalization of the proposed CSN techniques, we experiment on the AR database [40]. We use two conventional color spaces: RGB and XYZ, and three hybrid color spaces: XGB, YRB and ZRG. Each of these color spaces is normalized using the proposed two CSN techniques. All color spaces and their normalized versions are evaluated in the way as suggested in Section 4.1. The resulting recognition rates are presented in Table 7. The recognition rate of the same evaluation method on the grayscale images is also listed for comparison. The results in Table 7 show that the two CSN techniques (CSN-I and CSN-II) consistently improve the recognition performance of color spaces by an average of more than 3%. These results demonstrate again the effectiveness of the proposed two CSN techniques.

Comparing the results in Table 7 with those in Tables 2 and 4, we can find that the performance improvement caused by the two CSN techniques on the FRGC database appears more significant than that on the AR database. This fact indicates that the advantage of the proposed CSN techniques becomes more evident in uncontrolled illumination conditions than in controlled conditions, noticing that many images in the FRGC database were taken

**Table 6**

Recognition rates of 10 randomly generated color spaces before color space normalization and after color space normalization.

No. of generated color space	Before CSN	After CSN-I	After CSN-II
1	84.1	95.0	94.6
2	80.3	92.1	91.8
3	80.4	94.6	88.7
4	83.5	92.5	90.8
5	82.6	93.7	93.8
6	79.1	93.3	94.1
7	79.7	92.9	93.9
8	74.9	93.5	93.4
9	79.4	93.8	94.0
10	82.5	94.1	94.3
Average	<b>80.65</b>	<b>93.55</b>	<b>92.94</b>

**Table 7**

Recognition rates of color spaces before color space normalization and after color space normalization on the AR database.

Color space	Before CSN	After CSN-I	After CSN-II
RGB	80.0	83.5	84.3
XYZ	79.0	81.9	83.3
XGB	78.9	83.5	83.7
YRB	79.9	83.5	83.5
ZRG	79.6	83.7	83.1
Grayscale	78.2		

**Table 8**

Recognition rates of color spaces before color space normalization and after color space normalization using the LBP features on the AR database.

Color space	Before CSN	After CSN-I	After CSN-II
RGB	92.4	93.9	94.6
XYZ	93.8	94.5	94.4
XGB	93.2	94.0	95.4
YRB	93.2	94.9	94.9
ZRG	93.0	94.5	94.5
Grayscale	91.9	–	–

in uncontrolled illumination conditions while all images in the AR database were taken in controlled conditions. The controlled images have good image quality, while the uncontrolled images display poor image quality, such as low resolution of the face region and possible blurring. Color cues play a significantly more important role in uncontrolled image recognition than in controlled image recognition. The performance difference between on grayscale images and on color images in the FRGC database is dramatically larger than that in the AR database, as presented in Tables 1 and 7. This conclusion is completely consistent with Yip and Sinha’s conclusion, i.e., the contribution of color cues becomes evident when shape cues are degraded [20].

In addition, we also perform feature extraction by applying local binary patterns (LBP) method [42] to each color component images and concatenate the LBP features of the three component images into one augmented pattern vector. We then perform PCA+FLD based on the augmented pattern vector and achieve the results as shown in Table 8. These results indicate that the proposed CSN techniques are still effective for improving the performance of the RGB, XYZ and their hybrid color spaces. By comparing the results in Tables 7 and 8, we find that LBP can significantly enhance the performance of color spaces (or grayscale images) on the AR database.

4.4. Why can the CSN techniques improve recognition performance?

In this subsection, we try to address the problem of why the proposed CSN techniques can improve the face verification and recognition performance. We will show that the proposed CSN techniques can greatly reduce the correlation of the three color component images and thus can significantly enhance the discriminating power of color spaces. Here, we only show the analysis results on the FRGC database for conciseness, although the similar results are achieved on the AR database.

4.4.1. Correlation analysis of color components

Let  $\mathbf{x} = (x_1, x_2, x_3)^T$  be the color component vector in the original RGB color space and  $\mathbf{y} = (y_1, y_2, y_3)^T$  be the color component vector in the transformed color space. The transformation from  $\mathbf{x}$  to  $\mathbf{y}$  is

given by

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \mathbf{A}_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{25}$$

In the transformed color space, the correlation between the two color components  $y_i = \mathbf{A}_i\mathbf{x}$  and  $y_j = \mathbf{A}_j\mathbf{x}$  is

$$\text{Cov}(y_i, y_j) = E(y_i - E y_i)(y_j - E y_j) = \mathbf{A}_i \{E(\mathbf{x} - E\mathbf{x})(\mathbf{x} - E\mathbf{x})^T\} \mathbf{A}_j^T = \mathbf{A}_i \Sigma \mathbf{A}_j^T, \tag{26}$$

where  $\Sigma = E(\mathbf{x} - E\mathbf{x})(\mathbf{x} - E\mathbf{x})^T$  is the covariance matrix of the color component vectors in the original RGB color space.

Accordingly, the correlation coefficients between  $y_i$  and  $y_j$  is

$$\rho(y_i, y_j) = \frac{\mathbf{A}_i \Sigma \mathbf{A}_j^T}{\sqrt{\mathbf{A}_i \Sigma \mathbf{A}_i^T} \sqrt{\mathbf{A}_j \Sigma \mathbf{A}_j^T}} \tag{27}$$

Since the color space is usually three-dimensional, there are totally three correlation coefficients:  $\rho(y_1, y_2)$ ,  $\rho(y_2, y_3)$  and  $\rho(y_1, y_3)$ . The average absolute correlation coefficient is defined as follows:

$$\rho = (|\rho(y_1, y_2)| + |\rho(y_2, y_3)| + |\rho(y_1, y_3)|) / 3 \tag{28}$$

We use Eq. (28) to measure the correlation of three color components of a color space. The obtained average absolute correlation coefficients corresponding to all color spaces mentioned before are presented in Tables 9 and 10.

From Tables 9 and 10, we can see that the average absolute correlation coefficients are decreased after color space normalization, no matter which CSN technique is used. These results indicate that the proposed CSN techniques can greatly reduce the correlation between the three color components. The reduced correlation makes the discriminative information contained in the three color component images as mutually complementary as possible. Therefore, the concatenation of the three color component images can make use of the discriminative information from the three color channels. For the color spaces such as RGB, XYZ and their hybrid color spaces, the three color components are strongly correlated. As a result, the discriminative information from the three color component images is highly redundant. The concatenation of these three color component images cannot help much for improving the recognition performance.

In summary, the proposed CSN techniques can significantly reduce the correlation between the three color component images and thus can enhance the discriminating power of the concatenated color component images. This offers an intrinsic (or internal) reason for why the CSN techniques can improve the face verification and recognition performance. In the following subsection, we will use the total Fisher discriminant criterion value to measure the discriminating power of the concatenated color component images and show that the proposed CSN techniques can really enhance the discriminating power of color spaces.

**Table 9**

Average absolute correlation coefficient comparison: before color space normalization and after color space normalization.

Color space	Before CSN	After CSN-I	After CSN-II
RGB	0.8109	0.5850	0.7775
XYZ	0.8868	0.6210	0.4281
XGB	0.8756	0.5142	0.5378
YRB	0.8229	0.6938	0.6923
ZRG	0.8200	0.2964	0.1548

**Table 10**

Average absolute correlation coefficient comparison of 10 randomly generated color spaces: before color space normalization and after color space normalization.

No. of generated color space	Before CSN	After CSN-I	After CSN-II
1	0.9859	0.5518	0.6436
2	0.9781	0.6306	0.6339
3	0.9938	0.5481	0.7181
4	0.9503	0.7564	0.7526
5	0.9897	0.5676	0.6192
6	0.9976	0.6454	0.5817
7	0.9911	0.5891	0.5665
8	0.9989	0.2978	0.2110
9	0.9960	0.6919	0.6891
10	0.9877	0.6094	0.5857
Average	<b>0.9869</b>	<b>0.5888</b>	<b>0.6001</b>

**Table 11**

Discriminant criterion value (DCV) comparison: before color space normalization and after color space normalization.

Color space	Before CSN	After CSN-I	After CSN-II
RGB	344.79	391.67	394.08
XYZ	307.08	382.84	395.91
XGB	328.39	397.19	397.83
YRB	326.13	386.51	387.45
ZRG	344.86	397.98	409.94

#### 4.4.2. Discriminating power analysis of color spaces

The discriminating power of color images corresponding to a given color space can be characterized by the discriminant criterion value of the adopted evaluation method—the Fisher linear discriminant analysis (FLD) method. The FLD method applies the following Fisher criterion to determine its discriminant projection basis vectors [32]:

$$J(\boldsymbol{\phi}) = \frac{\boldsymbol{\phi}^T \mathbf{S}_b \boldsymbol{\phi}}{\boldsymbol{\phi}^T \mathbf{S}_w \boldsymbol{\phi}}, \quad (29)$$

The Fisher criterion represents the ratio of the between-class scatter to the within-class scatter of the data that are projected onto  $\boldsymbol{\phi}$ . Maximizing this criterion is equivalent to solving a generalized eigenvalue problem. The generalized eigenvector  $\boldsymbol{\phi}_1$  of  $\mathbf{S}_b \boldsymbol{\phi} = \lambda \mathbf{S}_w \boldsymbol{\phi}$  corresponding to the largest eigenvalue  $\lambda_1$  is chosen as a projection basis vector for two-class problems. The largest eigenvalue is the optimal value of the Fisher criterion, i.e.,  $J(\boldsymbol{\phi}_1) = \lambda_1$ , which determines the clustering performance of the data that are projected onto  $\boldsymbol{\phi}_1$ . Generally, for the given data, the larger the eigenvalue  $\lambda_1$  is, the better the clustering performance is. For multi-class problems, we usually choose a set of  $d$  eigenvectors  $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \dots, \boldsymbol{\phi}_d$  of  $\mathbf{S}_b \boldsymbol{\phi} = \lambda \mathbf{S}_w \boldsymbol{\phi}$  corresponding  $d$  largest eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_d$ . The sum of these  $d$  largest eigenvalues, i.e., the total Fisher discriminant criterion value (DCV) is computed as follows:

$$\text{DCV} = \sum_{j=1}^d \lambda_j \quad (30)$$

The DCV is thus used as a measure for evaluating the clustering performance (or the generalization performance of LDA) of the multi-class data that are projected onto  $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \dots, \boldsymbol{\phi}_d$ . Here, we use this measure to evaluate discriminating power of color images corresponding to a color space.

The total DCVs corresponding to all color spaces mentioned before are presented in Tables 11 and 12. From these tables, we can see that all DCVs are significantly increased after color space

**Table 12**

Discriminant criterion value (DCV) comparison of 10 randomly generated color spaces: before color space normalization and after color space normalization.

No. of generated color space	Before CSN	After CSN-I	After CSN-II
1	283.8168	414.9811	411.8312
2	270.5928	376.2171	363.7114
3	258.8294	409.3863	307.7064
4	286.9860	360.4551	344.2855
5	277.8274	373.4547	377.3002
6	244.3608	373.6426	371.5806
7	254.2313	394.5399	406.0598
8	235.5036	388.8860	389.1861
9	247.2688	394.2379	397.3350
10	271.9312	384.8894	384.8185
Average	<b>263.1348</b>	<b>387.0690</b>	<b>375.3815</b>

normalization, no matter which CSN technique is used. These results indicate that the proposed CSN techniques can greatly enhance the discriminating power of color spaces, which offers a direct reason for why the CSN techniques can improve the face verification and recognition performance.

## 5. Conclusions and discussions

This paper presents the concept of color space normalization (CSN) and two CSN techniques for enhancing the discriminating power of color spaces for face recognition. Our experimental results reveal that some color spaces, like RGB and XYZ, are relatively weak for recognition, whereas other color spaces, such as  $I_1 I_2 I_3$ , YUV, YIQ and LSLM, are relatively powerful. The proposed CSN techniques are applied to the RGB and XYZ color spaces, the three hybrid color spaces XGB, YRB and ZRG which are generated by configuring the components from the RGB and XYZ color spaces, and the 10 randomly generated color spaces. All experimental results demonstrated the effectiveness of the proposed CSN techniques.

To address the problem of why the CSN techniques can improve the face recognition performance of weak color spaces, we perform the correlation analysis on color component images corresponding to different color spaces and show that the proposed CSN techniques can significantly reduce the correlation between color component images and thus can enhance the discriminating power of the concatenated color component images. We further use the total Fisher discriminant criterion value to measure the discriminating power of the concatenated color component images and show that the proposed CSN techniques can really enhance the discriminating power of color spaces.

Finally, it should be pointed out that the focus of this paper is on validating the effectiveness of the color space normalization techniques for color images based face recognition. We only use a basic face feature extraction method, i.e., Fisher linear discriminant analysis, for performance evaluation. If using and combining more complicated feature extraction methods, such as kernel Fisher discriminant [31], Gabor wavelet [31], local binary patterns (LBP) [42], we can achieve state-of-the-art FRGC verification results based on the normalized color spaces. Additionally, we can further improve the verification rates of color spaces once the z-score normalization technique is applied [41].

## Acknowledgments

The authors would like to thank the anonymous reviewers for their critical and constructive comments and suggestions. This

work was partially supported by Award no. 2006-IJ-CX-K033 awarded by the National Institute of Justice, Office of Justice Programs, US Department of Justice. Dr. Yang was also supported by the National Science Foundation of China under Grant Nos. 60973098 and 60632050, the NUST Outstanding Scholar Supporting Program, and the Program for New Century Excellent Talents in University of China. Dr. Lei Zhang was supported by the Hong Kong RGC General Research Fund (PolyU 5351/08E).

## References

- [1] J. Luo, D. Crandall, Color object detection using spatial-color joint probability functions, *IEEE Transactions on Image Processing* 15 (6) (2006) 1443–1453.
- [2] R.L. Hsu, M. Abdel-Mottaleb, A.K. Jain, Face detection in color images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (5) (2002) 696–706.
- [3] O. Ikeda, Segmentation of faces in video footage using HSV color for face detection and image retrieval, in: *International Conference on Image Processing (ICIP 2003)*, 2003.
- [4] Y. Wu, T.S. Huang, Nonstationary color tracking for vision-based human-computer interaction, *IEEE Transactions on Neural Networks* 13 (4) (2002) 948–960.
- [5] T. Gevers, H. Stokman, Robust histogram construction from color invariants for object recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26 (1) (2004) 113–118.
- [6] A. Diplaros, T. Gevers, I. Patras, Combining color and shape information for illumination-viewpoint invariant object recognition, *IEEE Transactions on Image Processing* 15 (1) (2006) 1–11.
- [7] G. Dong, M. Xie, Color clustering and learning for image segmentation based on neural networks, *IEEE Transactions on Neural Networks* 16 (4) (2005) 925–936.
- [8] H.Y. Lee, H.K. Lee, Y.H. Ha, Spatial color descriptor for image retrieval and video segmentation, *IEEE Transactions on Multimedia* 5 (3) (2003) 358–367.
- [9] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, Content-based image retrieval at the end of the early years, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (12) (2000) 1349–1380.
- [10] M.J. Swain, D.H. Ballard, Color indexing, *International Journal of Computer Vision* 7 (1) (1991) 11–32.
- [11] B.V. Funt, G.D. Finlayson, Color constant color indexing, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (5) (1995) 522–529.
- [12] D.A. Adjeroh, M.C. Lee, On ratio-based color indexing, *IEEE Transactions on Image Processing* 10 (1) (2001) 36–48.
- [13] G. Healey, D.A. Slater, Global color constancy: recognition of objects by use of illumination invariant properties of color distributions, *Journal of the Optical Society of America A* 11 (11) (1994) 3003–3010.
- [14] G.D. Finlayson, S.D. Hordley, P.M. Hubel, Color by correlation: a simple, unifying framework for color constancy, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23 (11) (2001) 1209–1221.
- [15] H. Stokman, T. Gevers, Selection and fusion of color models for image feature detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (3) (2007) 371–381.
- [16] A.R. Weeks, *Fundamentals of Electronic Image Processing*, SPIE Optical Engineering Press, IEEE Press, Washington, USA, 1996.
- [17] W.H. Buchsbaum, *Color TV Servicing*, third ed., Prentice-Hall, NJ, Englewood Cliffs, 1975.
- [18] R. Kemp, G. Pike, P. White, A. Musselman, Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues, *Perception* 25 (1996) 37–52.
- [19] L. Torres, J.Y. Reutter, L. Lorente, The importance of the color information in face recognition, *International Conference on Image Processing (ICIP 99)* 3 (1999) 627–631.
- [20] A. Yip, P. Sinha, Role of color in face recognition, MIT tech report (ai.mit.com) AIM-2001-035 CBCL-212, 2001.
- [21] M. Rajapakse, J. Tan, J. Rajapakse, Color channel encoding with NMF for face recognition, *International Conference on Image Processing (ICIP '04)* 3 (2004) 2007–2010.
- [22] C. Xie, B.V.K. Kumar, Quaternion correlation filters for color face recognition, *Proceedings of the SPIE* 5681 (2005) 486–494.
- [23] P. Shih, C. Liu, Improving the face recognition grand challenge baseline performance using color configurations across color spaces, in: *IEEE International Conference on Image Processing, ICIP 2006*, 2006, October 8–11, Atlanta, GA.
- [24] P. Shih, C. Liu, Comparative assessment of content-based face image retrieval in different color spaces, *International Journal of Pattern Recognition and Artificial Intelligence* 19 (7) (2005) 873–893.
- [25] P. Shih, *Facial analysis in video: detection and recognition*, Ph.D. Dissertation, New Jersey Institute of Technology, 2006.
- [26] C. Jones III, A.L. Abbott, Color face recognition by hypercomplex gabor analysis, in: *Seventh International Conference on Automatic Face and Gesture Recognition (FGR 2006)*, April, 2006.
- [27] J. Yang, C. Liu, Color image discriminant models and algorithms for face recognition, *IEEE Transactions on Neural Networks* 19 (12) (2008) 2088–2098.
- [28] J. Yang, C. Liu, A. Discriminant color space method for face representation and verification on a large-scale database, in: *International Conference on Pattern Recognition 2008 (ICPR 2008)*, Tampa, Florida, USA, 2008.
- [29] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek, Overview of the face recognition grand challenge, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2005.
- [30] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, W. Worek, Preliminary face recognition grand challenge results, in: *Proceedings of the Seventh International Conference on Automatic Face and Gesture Recognition (FGR'06)*, 2006.
- [31] C. Liu, Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (5) (2006) 725–737.
- [32] Y. Ohta, T. Kanade, T. Sakai, Color information for region segmentation, *Computer Graphics and Image Processing* 13 (3) (1980) 222–241.
- [33] P. Colantoni, Color space transformations, <<http://www.couleur.org/>>.
- [34] D.L. Swets, J. Weng, Using discriminant eigenfeatures for image retrieval, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (8) (1996) 831–836.
- [35] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (7) (1997) 711–720.
- [36] W. Zhao, A. Krishnaswamy, R. Chellappa, D. Swets, J. Weng, Discriminant analysis of principal components for face recognition, in: H. Wechsler, P.J. Phillips, V. Bruce, F.F. Soulie, T.S. Huang (Eds.), *Face Recognition: From Theory to Applications*, Springer, 1998, pp. 73–85.
- [37] C. Liu, H. Wechsler, Robust coding schemes for indexing and retrieval from large face databases, *IEEE Transactions on Image Processing* 9 (1) (2000) 132–137.
- [38] J. Yang, J.Y. Yang, Why can LDA be performed in PCA transformed space?, *Pattern Recognition* 36 (2) (2003) 563–566.
- [39] A.R. Webb, *Statistical pattern recognition*, second ed., Wiley, New York, 2002.
- [40] A.M. Martinez, R. Benavente, The AR face database, CVC Technical Report #24, June 1998.
- [41] J. Yang, C. Liu, Horizontal and vertical 2DPCA-based discriminant analysis for face verification on a large-scale database, *IEEE Transactions on Information Forensics and Security* 2 (4) (2007) 781–792.
- [42] T. Ahonen, A. Hadid, M. Pietikainen, Face description with local binary patterns: application to face recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (12) (2006).
- [43] P. Shih, C. Liu, Evolving effective color features for improving FRGC baseline performance, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 25–25 (2005) 156–256.
- [44] Z. Liu, C. Liu, A hybrid color and frequency features method for face recognition, *IEEE Transactions on Image Processing* 17 (10) (2008) 1975–1980.
- [45] J.Y. Choi, Y.M. Ro, K.N. Plataniotis, Color face recognition for degraded face images, *IEEE Transactions on Systems, Man, and Cybernetics. Part B* 39 (5) (2009) 1217–1230.

**About the Author**—JIAN YANG received the BS degree in mathematics from the Xuzhou Normal University in 1995. He received the MS degree in applied mathematics from the Changsha Railway University in 1998 and the Ph.D. degree from the Nanjing University of Science and Technology (NUST), on the subject of pattern recognition and intelligence systems in 2002. In 2003, he was a postdoctoral researcher at the University of Zaragoza, and in the same year, he was awarded the RYC program Research Fellowship sponsored by the Spanish Ministry of Science and Technology. From 2004 to 2006, he was a postdoctoral fellow at Biometrics Centre of Hong Kong Polytechnic University. From 2006 to 2007, he was a postdoctoral fellow at Department of Computer Science of New Jersey Institute of Technology. Now, he is a professor in the School of Computer Science and Technology of NUST. He is the author of more than 50 scientific papers in pattern recognition and computer vision. His research interests include pattern recognition, computer vision and machine learning. Currently, he is an associate editor of *Pattern Recognition Letters* and *Neurocomputing*, respectively.

**About the Author**—CHENGJUN LIU received the Ph.D. from George Mason University in 1999, and he is presently an associate professor of computer science at New Jersey Institute of Technology. His research interests are in computer vision (face/iris detection and recognition, video processing), machine learning (statistical learning, kernel methods, similarity measures), security (biometrics), pattern recognition, and image processing. His recent work focuses on biometric research (with an emphasis on face

and iris recognition) and biometric system development. The class of new methods he has developed includes the Bayesian discriminating features method (BDF), the probabilistic reasoning models (PRM), the enhanced fisher models (EFM), the enhanced independent component analysis (EICA), the shape and texture-based Fisher method (STF), the Gabor–Fisher classifier (GFC), and the independent Gabor features (IGF) method. He has also pursued the development of novel evolutionary methods leading to the development of the evolutionary pursuit (EP) method for pattern recognition in general, and face recognition in particular.

**About the Author**—LEI ZHANG received the B.S. degree in 1995 from Shenyang Institute of Aeronautical Engineering, Shenyang, P.R. China, the M.S. and Ph.D. degrees in automatic control theory and engineering from Northwestern Polytechnical University, Xi'an, P.R. China, respectively, in 1998 and 2001. From 2001 to 2002, he was a research associate in the Department of Computing, The Hong Kong Polytechnic University. From January 2003 to January 2006 he worked as a postdoctoral fellow in the Department of Electrical and Computer Engineering, McMaster University, Canada. Since January 2006, he has been an assistant professor in the Department of Computing, The Hong Kong Polytechnic University. His research interests include image and video processing, biometrics, pattern recognition, multisensor data fusion and optimal estimation theory, etc.