Pattern Recognition 83 (2018) 456–468

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

# Color space construction by optimizing luminance and chrominance components for face recognition

Ze Lu*, Xudong Jiang, Alex Kot

*Nanyang Technological University, 50 Nanyang Drive, 637553, Singapore*

ARTICLE INFO

*Article history:*
Received 19 June 2017
Revised 2 May 2018
Accepted 19 June 2018
Available online 20 June 2018

*Keywords:*
Color face recognition
Color space
Color sensor analysis
Chrominance subspace
Discriminant analysis
Covariance analysis

ABSTRACT

In recent face recognition techniques utilizing the color information, researchers tried to select one conventional color space or learn a color space from the given training data to achieve better performance. RQCr, DCS and ZRG-NII color spaces have gained reputation as effective color spaces in which the face recognition performs better than in the others. However, at the moment, how to construct effective color spaces for face recognition has not been thoroughly studied. In this paper, we propose a color space $LuC_1C_2$ based on a framework of constructing effective color spaces for face recognition tasks. It is composed of one luminance component $Lu$ and two chrominance components $C_1$, $C_2$. The luminance component $Lu$ is selected from 4 different luminance candidates by comparing their R,G,B coefficients and color sensor properties. For the two chrominance components $C_1$, $C_2$, the directions of their transform vectors are determined by the discriminant analysis and the covariance analysis in the chrominance subspace of the RGB color space. The magnitudes of their transform vectors are determined by the discriminant values of $Lu$, $C_1$, $C_2$. Extensive experiments are conducted on 4 benchmark databases to evaluate our proposed color space $LuC_1C_2$. The experimental results obtained by using 2 different color features and 3 different dimension reduction methods show that our proposed color space $LuC_1C_2$ achieves consistently better face recognition performance than state-of-the-art color spaces on 3 databases. We also show that the proposed color space achieves higher face verification rate than the state of the arts on FRGC database. Furthermore, the face verification performance is improved significantly by combining CNN features with simple raw-pixel features from the proposed $LuC_1C_2$ color space on both LFW and FRGC databases.



## 1. Introduction

Recent research efforts reveal that color provides useful information for machine learning and pattern recognition tasks, such as color constancy [1,2] and color histograms [3,4]. Color in the machine vision system is defined by a combination of 3 color components specified by a color space. For visual recognition tasks like object detection, retrieval and recognition, different color spaces possess significantly different characteristics and effectiveness in terms of discriminating power [5]. For instance, the hue, saturation, value (HSV) color space and the luminance, chrominance-blue, chrominance-red (YCbCr) color space have been used for face detection by many researchers [6,7]. And the component image R in the RGB color space has been shown to be more effective for face recognition than the component images in several other color spaces [8].

Face recognition (FR) has become a very active research area driven mainly by its broad applications in human-computer interaction, homeland security, and entertainment [9–11]. Color information plays a discriminative and complementary role in face recognition tasks. In [12], the recognition performance of color invariants on a large database with extreme illumination variations suggests that the use of color information may significantly improve the greyscale-based matching algorithms. Torres et al. [13] applied a modified PCA scheme to face recognition and their results show that the use of color information improves the recognition rate when compared to the same scheme using the luminance information only. The improvement can be significant when large facial expression and illumination variations are present or the resolution of face images is low [14,15]. Since then, considerable research efforts have been devoted to the efficient utilization of facial color information for enhancing the face recognition performance, such as the fusion of color, local spatial and global frequency information in [16], the color space normalization in [17], the color channel fusion in [18] and the preprocessing mismatch effects in color images in [19].

* Corresponding author.
*E-mail addresses:* zlu008@ntu.edu.sg (Z. Lu), exdjiang@ntu.edu.sg (X. Jiang), eackot@ntu.edu.sg (A. Kot).

https://doi.org/10.1016/j.patcog.2018.06.015
0031-3203/

Many color spaces have been proposed to find the optimal way of representing color images for face recognition. These color spaces are usually defined by transformations of the RGB color space. In the early studies, color spaces were constructed through a combination of intuition and empirical comparisons without a systematic strategy, such as YCbCr, YIQ [8], YUV [13] and the hybrid color space YQCr in [20], where the Y and Q color components are from the YIQ color space and the Cr color component is from the YCbCr color space. Among different color space configurations discussed in [15], RQCr, where R is taken from the RGB color space and Q, Cr are taken from the YIQ and YCbCr color spaces respectively, shows the best face recognition performance on more than 3000 color facial images collected from three standard face databases [15]. Moreover, the R channel is known to be the best monochrome channel and QCr is the best chromaticity-component combination for achieving high face recognition performance [21,22]. And the RQCr color space was used in [16] to extract multiple color features from face images.

Rather than selecting color components heuristicly from a pool of color spaces, the DCS color space was proposed by seeking 3 sets of optimal coefficients to linearly combine the R, G and B components based on a discriminant criterion [23]. The experimental results show that the DCS color space is effective in enhancing the face recognition performance of RGB and Ig(r-g) color spaces. Compared with other learning-based color spaces such as ICS and UCS [24], DCS retains the spatial structure information of face images. In [17], a common characteristic of effective color spaces for face recognition was found out by analyzing their transformation matrices from the RGB color space. Based on this characteristic of effective color spaces, the authors proposed two color space normalization techniques, which are able to convert weak color spaces into effective ones, so that better face recognition performance can be achieved by using these normalized color spaces. Among normalized color spaces evaluated in [17], the ZRG-NII color space achieved the best FR performance. Therefore in [25], RQCr and ZRG-NII were considered to be the most two effective color spaces for face recognition.

RQCr, DCS and ZRG-NII color spaces do show better face recognition performance than the others on some databases. However, their performance is not consistent on different databases. Besides, they were proposed based on different criteria. In this paper, we propose a framework to construct a color space for obtaining high face recognition accuracy. By analysing color spaces that demonstrate good classification capabilities, we find that they are composed of one luminance component and two chrominance components except the learning-based DCS color space. This configuration reduces the correlation of the three color components and thus enhances the discriminating power of the whole color space. Based on the proposed framework, we construct a color space $LuC_1C_2$. The luminance component $Lu$ is chosen among four luminance candidates from existing color models by investigating their R,G,B coefficients and the color sensor properties. For the two chrominance components $C_1$, $C_2$, the directions of their transform vectors are derived by the discriminant analysis and the covariance analysis in the chrominance subspace of the RGB color space. The magnitudes of their transform vectors are derived according to the discriminant values of $Lu$, $C_1$, $C_2$. In the experimental part, we firstly validate the dependence of the face recognition performance on the correlation of two chrominance components. Then the FR performance of the proposed $LuC_1C_2$ color space is compared with those of DCS, RQCr, ZRG-NII and RGB color spaces on 3 benchmark databases (AR, GT and FRGC) using 2 distinct color features and 3 different dimension reduction methods. After that, we show that the fusion of multiple features extracted from the proposed color space achieves higher face verification rate than recently published state-of-the-art results on FRGC database. Finally on LFW and FRGC databases, we show that features from our proposed color space $LuC_1C_2$ can be used to enhance the face verification performance of both pre-trained and fine-tuned CNN models.

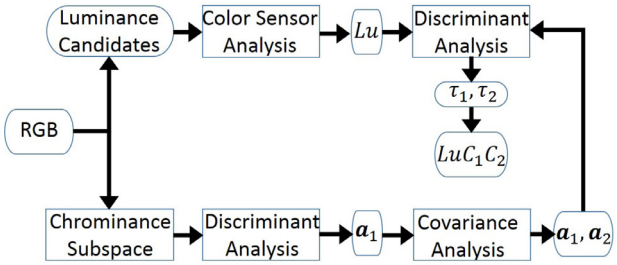



**Fig. 1.** The framework of constructing an effective color space for face recognition. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The novelty of this paper comes from (1) the framework of constructing effective color spaces for face recognition tasks; (2) R,G,B coefficient analysis and color sensor analysis for luminance components; (3) the computation of two chrominance components; (4) the color space $LuC_1C_2$ that consistently performs better than state-of-the-art color spaces on three benchmark databases using two distinct features and three different dimension reduction methods; (5) achieving the state-of-the-art face verification rate on FRGC database using the proposed color space; 6) improving the face verification performance of CNN and traditional features by using features from the proposed color space $LuC_1C_2$ on LFW and FRGC databases.

## 2. The proposed color space

In this section, the framework of constructing a color space for achieving high face recognition performance is presented first and then the $LuC_1C_2$ color space is introduced.

### 2.1. Overview of the proposed approach

In video, the luminance (also known as gray, brightness, intensity, or lightness) is formed as a weighted sum of R,G,B components to represent the brightness of pixels, and the chrominance conveys the color information, separately from the accompanying luminance [26]. Luminance and chrominance are usually processed separatedly in video compression standards such as MPEG and JPEG [27]. Motivated by this, we study the framework of constructing well-performing color spaces for FR tasks. To begin with, several color spaces which have shown good classification capabilities for face recognition are investigated. These color spaces include $I_1I_2I_3$ [28], YUV, YIQ, YCbCr, [13,29], YQCr [20], RQCr [15], LSLM [30] and ZRG-NII [17]. One common characteristic of these color spaces is that they are all composed of one luminance component ($I_1$, Y, R, L, Z) and two chrominance components ($I_2I_3$, UV, IQ, CbCr, QCr, SLM, $N_1N_2$). Experimental results in [31] suggest that color cues do play a role in face recognition and their contribution becomes evident when shape cues are degraded. The separation of luminance and chrominance makes the correlation of three color components reduced and the discriminative information contained in the three color components become mutually complementary [17]. Based on these analyses, we propose the framework of constructing an effective color space for face recognition on Fig. 1.

In the proposed framework, three color components of the proposed color space, $Lu$, $C_1$, $C_2$, are derived by linear transformations of the RGB color space. The luminance component $Lu$ is selected from 4 different luminance candidates. To obtain the two chrominance components $C_1$, $C_2$, we first calculate the directions $\mathbf{a}_1$, $\mathbf{a}_2$ of

their transform vectors by the discriminant analysis and the covariance analysis in the chrominance subspace of the RGB color space, and then calculate the magnitudes $\tau_1$, $\tau_2$ of their transform vectors according to the discriminant values of *Lu*, $C_1$, $C_2$.

### 2.2. Selection of the luminance component

Color spaces are usually defined by transformations of the RGB color space. The transformations involved are either linear or non-linear. As mentioned in [17], current research findings show that some linear color spaces, which are derived by linear transformations from the RGB color space, perform much better than those derived by non-linear transformations from the RGB color space. Thus in many face recognition algorithms, an image in the RGB color space was firstly converted into a monochrome image by linearly combining its three color components [32]. However, theoretical and experimental justifications are lacking for investigating which monochrome image is the best representation of the color image for the face recognition purpose. In this section, 4 widely-used luminance components including $I_1$ from $I_1I_2I_3$ [33], R from RGB, Y from YUV [29] and L from LSLM [30] are analysed. They are computed from the RGB color space by

$$\begin{pmatrix} I_1 \\ R \\ Y \\ L \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1 & 0 & 0 \\ 0.299 & 0.587 & 0.114 \\ 0.209 & 0.715 & 0.076 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \tag{1}$$

The $I_1I_2I_3$ color space proposed by Ohta et al. uses a K-L transformation to decorrelate the R,G,B components. The effectiveness of the luminance component $I_1$ for face verification was shown in [28]. But $I_1$ implicitly assumes a uniform distribution of light energy over the entire color space.

The color component R in the RGB color space has been shown to perform better than other intensity images including $I_1$ and Y/Gray for face retrieval in [8]. In addition, the R channel of skin-tone color is known to be the best monochrome channel for face recognition [21]. However, the performance superiority of the R component over the others for face recognition is reflected only on the FRGC database [34]. In general, it is not effective to select R as the luminance component as R discards useful information by ignoring G and B components.

YUV, YIQ, YCbCr are three color spaces commonly used for efficient video transmission. In these 3 color spaces, the R,G,B components are transformed into the luminance component Y and two chrominance components. Y performs the best to display pictures on monochrome (black and white) televisions [35]. And it has been used in a great number of traditional face recognition algorithms [36–38]. Nevertheless, there is no proof that Y is the optimal monochromatic form of color images for face recognition tasks. The LSLM color space is obtained from the RGB color space by a linear transformation based on the opponent signals of the cones: blackwhite, redgreen, and yellowblue. And the L component describes the luminance information.

Fig. 2[39] shows the normalized response of human cone cells against the wavelength of light, where the shapes of the curves are obtained by measurement of the light absorption by the cones, and the relative heights for the three types (L, S, M) are set equal. Comparing red, green and blue lights, we can see that the human eye is most sensitive to green light, followed by red light and then blue light. The blue light mainly stimulates S cones, the red light stimulates the two most common (M and L) of the three kinds of cones, and the green light stimulates M and L cones at a higher response level than the red light. Moreover, noise in the green channel is much less than that in the other two primary colors [40].

To produce a color image in the machine vision system, a digital device needs three different kinds of sensors to acquire the red, green, and blue parts of the spectrum [42]. 50% of the camera sensors in the color filter array are green sensors, which are twice as many as red or blue sensor as shown on Fig. 3(a). Fig. 3(b) shows the spectral sensitivity of the Nikon-D70 sensors for different wavelengths of the light. One observation from Fig. 3(b) is that the highest sensitivity level of G sensor is almost twice as high as that of R or B sensor. Another obervation from Fig. 3(b) is that, above the sensitivity level of 0.1, the passband width of G sensor (150 nm) is much larger than that of R sensor (90 nm) or B sensor (100 nm). What's more, the spectrum of G sensor spreads from the central to the two sides of the visible spectrum rather than reside at one side of the visible spectrum like those of R and B sensors. To summarize, 50% of camera sensors in the color filter array are green sensors, each of them has higher sensitivity level and broader passband compared with those of red and blue sensors. Therefore, it is more reasonable to assign the largest weight (significantly larger than 0.5) to G in the linear combination of R, G, B components for the luminance component *Lu*. The second largest weight should be assigned to R since it performs much better than B for face recognition in [8,15]. Following these analyses, we select L that has R, G, B weights of 0.209, 0.715 and 0.076, respectively, as the luminance component *Lu*.

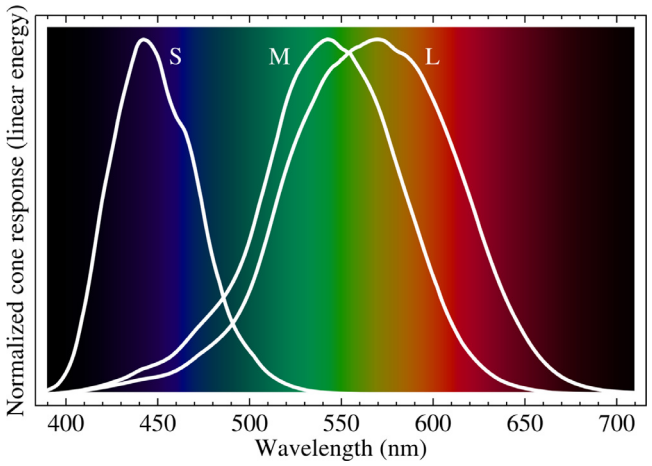



**Fig. 2.** The normalized response of human cone cells for different wavelengths of light. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

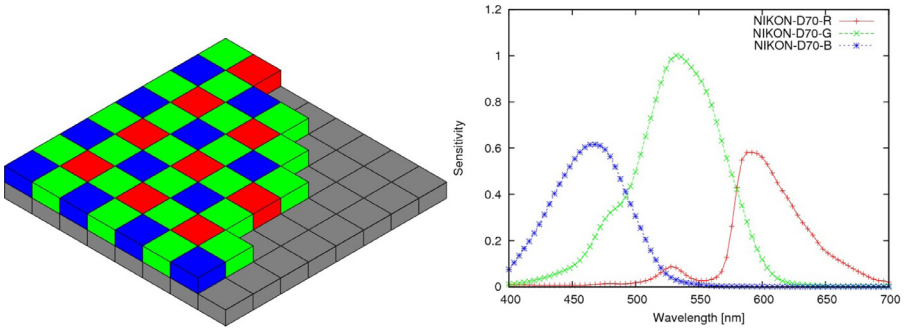
### 2.3. Extraction of the two chrominance components

Chrominance is used in video systems to convey the color information of the picture, separately from the accompanying luminance signal [26]. Different chrominance components are computed from the RGB color space by linear transformations as shown below:

$$\begin{pmatrix} I_2 \\ I_3 \\ U \\ V \\ I \\ Q \\ S \\ LM \\ N1 \\ N2 \end{pmatrix} = \begin{pmatrix} 0.5 & 0 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.1471 & -0.2888 & 0.4359 \\ 0.6148 & -0.5148 & -0.1 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \\ 0.209 & 0.715 & -0.924 \\ 3.148 & -2.799 & -0.349 \\ 0.821 & -0.215 & -0.606 \\ -0.179 & 0.785 & -0.606 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \tag{2}$$

These chrominance components are from the $I_1I_2I_3$, YUV, YIQ, LSLM, and ZRG-NII color spaces respectively.

It is easy to find a common characteristic of these chrominance components that the sum of R,G,B coefficients $\mathbf{u^T} = (u_1, u_2, u_3)$ is




(a) (b)

**Fig. 3.** (a) The color filter array in cameras and (b) the camera spectral sensitivity of nikon-D70 [41]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

zero for every chrominance component C as shown in (3) and (4).

$$C = \mathbf{u}^T \begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \tag{3}$$

$$\sum_{j=1}^{3} u_j = 0. \tag{4}$$

Our explanation for this characteristic is that, the luminance contained in R, G and B components gets cancelled out and the remaining information forms a chrominance component C. In order to generate a chrominance component, the sum of $u_1$, $u_2$, $u_3$ needs to be zero. Therefore, $\mathbf{u}$ is orthogonal to $\mathbf{b} = (1, 1, 1)^T$ in the 3-dimensional RGB color space as follows:

$$\sum_{j=1}^{3} u_j = \mathbf{u}^T \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \mathbf{u}^T \mathbf{b} = 0. \tag{5}$$

Now the question is: how to find two appropriate $3 \times 1$ vectors $\mathbf{u}_1$, $\mathbf{u}_2$ to linearly combine R,G,B in (3) and produce two chrominance components $C_1$, $C_2$ respectively. According to (5), $C_1$, $C_2$ will be the two chrominance components as long as $\mathbf{u}_1$ and $\mathbf{u}_2$ lie in the 2-dimensional subspace orthogonal to $\mathbf{b}$, which is referred to as the chrominance subspace in our paper.

But which pair of $\mathbf{u}_1$, $\mathbf{u}_2$ lying in the 2-dimensional subspace orthogonal to $\mathbf{b}$ is the optimum choice for achieving high face recognition performance is still a question. It has been shown in recent studies [43–45] that the face recognition performance is related with both the discriminating power and the correlation of color components. To find $\mathbf{u}_1$, $\mathbf{u}_2$ that maximize the discriminative power and minimize the correlation of $C_1$, $C_2$, we apply the discriminant analysis and the covariance analysis to $\mathbf{u}_1$, $\mathbf{u}_2$ in the chrominance subspace of the RGB color space. As the eigenvector derived by discriminant analysis and the vector derived by covariance analysis are of arbitrary lengths, they provide only the direction information. Therefore, we represent $\mathbf{u}_1$, $\mathbf{u}_2$ by the magnitude and the direction as follows:

$$\mathbf{u}_1 = \tau_1 \mathbf{a}_1, \mathbf{u}_2 = \tau_2 \mathbf{a}_2, \tag{6}$$

where $\tau_i, \mathbf{a}_i (i = 1, 2)$ indicate the magnitude and the direction of vector $\mathbf{u}_i$ respectively.

The Fisher criterion [23] as defined in (7) is an effective criterion for classification tasks because it maximizes the between-class variation and minimizes the within-class variation of the projected data simultaneously.

$$\max_{\mathbf{x}} \mathbf{J}(\mathbf{x}) = \max_{\mathbf{x}} \frac{\mathbf{x}^T \mathbf{S}_b \mathbf{x}}{\mathbf{x}^T \mathbf{S}_w \mathbf{x}} \tag{7}$$

In (7), $\mathbf{J}$ represents the ratio of the between-class variation to the within-class variation of the data projected onto $\mathbf{x}$. $\mathbf{S}_b$, $\mathbf{S}_w$ are the between-class scatter matrix and the within-class scatter matrix of the input data, respectively.

In our case, $C_1$ will be a chrominance component as long as $\mathbf{u}_1$ lies in the 2-dimensional chrominance subspace (5). Under this condition, the Fisher criterion can be utilized to find the optimal $\mathbf{u}_1$ which maximizes the discriminative power of $C_1$. Pixel values of color images are transformed from the 3-dimensional RGB color space to the 2-dimensional chrominance subspace as below:

$$\mathbf{B} = \begin{pmatrix} \mathbf{n}_1^T \\ \mathbf{n}_2^T \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}, \tag{8}$$

where $\mathbf{n}_1$, $\mathbf{n}_2$ are two basis vectors that span the 2-dimensional chrominance subspace $\mathbf{B}$ orthogonal to $\mathbf{b}$. Although $\mathbf{n}_1$, $\mathbf{n}_2$ are not unique, the final results are independent of the choice of $\mathbf{n}_1$, $\mathbf{n}_2$, because any pair of orthogonal vectors of unit length orthogonal to $\mathbf{b}$ determines the same subspace $\mathbf{B}$. Suppose each image in the training dataset contains k pixels. Let $\mathbf{B}_{ij} \in \mathbf{R}^{2 \times k}$ denote the projection of the $j$th RGB image of class $i$ into the 2-dimensional chrominance subspace $\mathbf{B}$, where $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, q_i$. $p$ indicates the number of classes and $q_i$ indicates the number of images for class $i$.

To compute the value of $\mathbf{J}$ in (7), we define the within-class scatter matrix $\mathbf{S}_w \in \mathbf{R}^{2 \times 2}$ and the between-class scatter matrix $\mathbf{S}_b \in \mathbf{R}^{2 \times 2}$ as follows:

$$\mathbf{S}_w = \sum_{i=1}^{p} \sum_{j=1}^{q_i} (\mathbf{B}_{ij} - \overline{\mathbf{B}}_i)(\mathbf{B}_{ij} - \overline{\mathbf{B}}_i)^T, \tag{9}$$

$$\mathbf{S}_b = \sum_{i=1}^{p} q_i (\overline{\mathbf{B}}_i - \overline{\mathbf{B}})(\overline{\mathbf{B}}_i - \overline{\mathbf{B}})^T, \tag{10}$$

where

$$\overline{\mathbf{B}}_i = \frac{1}{q_i} \sum_{j=1}^{q_i} \mathbf{B}_{ij}, \tag{11}$$

$$\overline{\mathbf{B}} = \frac{1}{q} \sum_{i=1}^{p} \sum_{j=1}^{q_i} \mathbf{B}_{ij}. \tag{12}$$

$\overline{\mathbf{B}}_i$ indicates the mean matrix of training samples in class $i$ and $\overline{\mathbf{B}}$ indicates the mean matrix of all training samples. $q$ is the number of all training samples and $q = \sum_{i=1}^{p} q_i$.

As both $\mathbf{S}_b$ and $\mathbf{S}_w$ are positive definite matrices, $\mathbf{J(x)}$ in (7) is actually a generalized Rayleigh quotient [46]. Thus maximizing $\mathbf{J(x)}$ is equivalent to solving a generalized eigenvalue problem

$$\mathbf{S}_b\mathbf{x} = \lambda\mathbf{S}_w\mathbf{x}, \tag{13}$$

where $\mathbf{x} \in \mathbf{R}^{2\times 1}$ is an eigenvector and $\lambda$ is the corresponding eigenvalue. After $\mathbf{S}_b$, $\mathbf{S}_w$ are substituted into (13), the derived eigenvector $\mathbf{x}$ corresponding to the largest eigenvalue is chosen as the optimal projection vector since the eigenvalue indicates the discriminative power of data projected to the corresponding eigenvector. Note that the eigenvector $\mathbf{x}$ obtained from (13) is of arbitrary length, which means it provides only the direction information. Here, we take $\mathbf{x}$ of the unit length. Then $\mathbf{a}_1$ is computed by

$$\mathbf{a}_1 = \begin{pmatrix} \mathbf{n}_1, \mathbf{n}_2 \end{pmatrix}\mathbf{x}. \tag{14}$$

And $C_1$ is given by

$$C_1 = \mathbf{u}_1^T\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \tau_1\mathbf{a}_1^T\begin{pmatrix} R \\ G \\ B \end{pmatrix}. \tag{15}$$

The key of color face recognition techniques is how to effectively utilize the complementary information between color components and remove their redundancy according to [43,47,48]. The dependence of face recognition performance on the correlation between two chrominance components $C_1$, $C_2$ is validated in our experimental part. The discriminative power of $C_1$ has been maximized in the previous subsection, we then derive $\mathbf{a}_2$ by minimizing the correlation between $C_1$, $C_2$. Similar to $C_1$, $C_2$ is given by

$$C_2 = \mathbf{u}_2^T\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \tau_2\mathbf{a}_2^T\begin{pmatrix} R \\ G \\ B \end{pmatrix}. \tag{16}$$

The covariance provides a measure of the strength of the correlation between two or more sets of random variates [49]. Hence the covariance between $C_1$ and $C_2$ is defined as below to measure the correlation between them.

$$\begin{aligned} &cov(C_1, C_2) \\ &= E((C_1 - E(C_1))(C_2 - E(C_2))^T) \\ &= \mathbf{u}_1^T E\left(\begin{pmatrix} R-\overline{R} \\ G-\overline{G} \\ B-\overline{B} \end{pmatrix}\begin{pmatrix} R-\overline{R} \\ G-\overline{G} \\ B-\overline{B} \end{pmatrix}^T\right)\mathbf{u}_2 \\ &= \tau_1\mathbf{a}_1^T\mathbf{M}\tau_2\mathbf{a}_2, \end{aligned} \tag{17}$$

where $E$ indicates the expectation and $\mathbf{M}$ is

$$\mathbf{M} = E\left(\begin{pmatrix} R-\overline{R} \\ G-\overline{G} \\ B-\overline{B} \end{pmatrix}\begin{pmatrix} R-\overline{R} \\ G-\overline{G} \\ B-\overline{B} \end{pmatrix}^T\right), \tag{18}$$

which can be estimated from the training data.

When $cov(C_1, C_2)$ equals zero, $C_1$, $C_2$ will be uncorrelated and the information contained in them will become most complementary to each other. So we make $cov(C_1, C_2)$ equal to zero to minimize the correlation of $C_1$, $C_2$:

$$cov(C_1, C_2) = \tau_1\mathbf{a}_1^T\mathbf{M}\tau_2\mathbf{a}_2 = 0. \tag{19}$$

$\mathbf{a}_2$ can be represented by the two basis vectors $\mathbf{n}_1$, $\mathbf{n}_2$ and its angle $\theta$ in the 2-dimensional chrominance subspace $\mathbf{B}$ as below:

$$\mathbf{a}_2 = cos(\theta)\mathbf{n}_1 + sin(\theta)\mathbf{n}_2, \theta \in [0, 2\pi). \tag{20}$$

Substitute (20) into (19), we have

$$\theta = tan^{-1}\left(-\frac{\mathbf{a}_1^T\mathbf{M}\mathbf{n}_1}{\mathbf{a}_1^T\mathbf{M}\mathbf{n}_2}\right). \tag{21}$$

With $\mathbf{a}_1$ determined in (14), $\mathbf{a}_2$ can be computed by (20) and (21).

*Lu*, $C_1$, $C_2$ possess different discriminating power for face recognition. The discriminating power of a color component can be characterized by the discriminant value. Suppose a column vector $\mathbf{s}_{ij}$ contains all pixel values of *Lu* of the $j$th image in class $i$, where $i = 1, 2, \ldots, p$ and $j = 1, 2, \ldots, q_i$. $p$ indicates the number of classes and $q_i$ indicates the number of images for class $i$. The discriminant value of *Lu*, $J_{Lu}$, is defined as follows:

$$v_w = \sum_{i=1}^{p}\sum_{j=1}^{q_i}(\mathbf{s}_{ij} - \overline{\mathbf{s}_i})^T(\mathbf{s}_{ij} - \overline{\mathbf{s}_i}), \tag{22}$$

$$v_b = q_i\sum_{i=1}^{p}(\overline{\mathbf{s}_i} - \overline{\mathbf{s}})^T(\overline{\mathbf{s}_i} - \overline{\mathbf{s}}), \tag{23}$$

$$J_{Lu} = \sqrt{\frac{v_b}{v_w}}, \tag{24}$$

where $\overline{\mathbf{s}_i}$ is the mean vector of training samples of class $i$ and $\overline{\mathbf{s}}$ is the mean vector of all training samples. The discriminant values of training images projected onto vectors $\mathbf{a}_1$ and $\mathbf{a}_2$, $J_{\mathbf{a}_1}, J_{\mathbf{a}_2}$, can be computed similarly to (22)–(24).

To represent $C_1$, $C_2$ by $\mathbf{u}_1$, $\mathbf{u}_2$ according to their discriminant values, we assign $\mathbf{u}_1$, $\mathbf{u}_2$ magnitudes $\tau_1$, $\tau_2$ proportional to the discriminant values of $C_1$, $C_2$ by the equation below:

$$\tau_0 : \tau_1 : \tau_2 = J_{Lu} : J_{\mathbf{a}_1} : J_{\mathbf{a}_2}, \tag{25}$$

where $\tau_0$ is the magnitude of the transform vector of *Lu*. Thus

$$\tau_1 = \frac{J_{\mathbf{a}_1}}{J_{Lu}}\tau_0, \tau_2 = \frac{J_{\mathbf{a}_2}}{J_{Lu}}\tau_0. \tag{26}$$

Two chrominance components $C_1$, $C_2$ are computed by

$$C_1 = \mathbf{u}_1^T\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \frac{J_{\mathbf{a}_1}}{J_{Lu}}\tau_0\mathbf{a}_1^T\begin{pmatrix} R \\ G \\ B \end{pmatrix}, \tag{27}$$

$$C_2 = \mathbf{u}_2^T\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \frac{J_{\mathbf{a}_2}}{J_{Lu}}\tau_0\mathbf{a}_2^T\begin{pmatrix} R \\ G \\ B \end{pmatrix}. \tag{28}$$

The extraction of two chrominance components $C_1$, $C_2$ from the RGB color space is summarized in **Algorithm 1**.

**Algorithm 1** Extraction of two chrominance components $C_1$, $C_2$.

1: Calculate two basis vectors $\mathbf{n}_1, \mathbf{n}_2$ of the 2-dimensional chrominance subspace orthogonal to $\mathbf{b}$ in (5).
2: Transform pixel values of color images from the 3-dimensional RGB color space to the 2-dimensional chrominance subspace $\mathbf{B}$ in (8) and compute the within-class scatter matrix $\mathbf{S}_w$ and the between-class scatter matrix $\mathbf{S}_b$ in (9)-(12).
3: Solve the eigenvalue problem in (13) to get $\mathbf{x}$ and calculate $\mathbf{a}_1$ in (14).
4: Compute $\mathbf{M}$ in (18) from training images and compute $\mathbf{a}_2$ by (20), (21).
5: Compute $J_{Lu}, J_{\mathbf{a}_1}, J_{\mathbf{a}_2}$ using (22)-(24) and substitute them into (27),(28) to compute $C_1, C_2$.

## 3. Experiments

The effectiveness of our proposed $LuC_1C_2$ color space for face recognition is extensively evaluated on four benchmark databases, AR [50], Georgia Tech (GT) [51], FRGC [34] and the LFW [52].

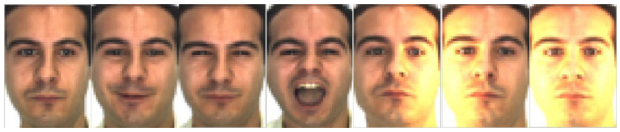


**Fig. 4.** Cropped images of the AR database.

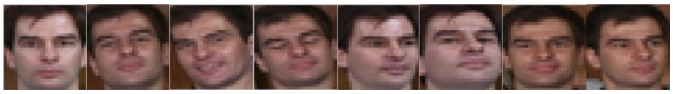


**Fig. 5.** Cropped images of the GT database.

To begin with, we validate the dependence of FR performance on the correlation of $C_1$, $C_2$ through experimentation. The evaluation of the proposed color space for face recognition consists of 3 parts. In the first part, we compare the face recognition performance of the proposed $LuC_1C_2$ color space with that of three state-of-the-art color spaces, DCS [23], RQCr [15] and ZRG-NII [17], and the fundamental color space, RGB. Their face recognition/verification rates are compared under various conditions: 2 different feature representations and 3 different dimension reduction methods on 3 databases. In the second part, we compare the face verification rate (FVR) of $LuC_1C_2$ with the state of the arts on FRGC database. Multiple features are extracted from $LuC_1C_2$ images and fused at the decision level. Finally, we use features from $LuC_1C_2$ color space to enhance the face verification performance of CNN and traditional features on LFW and FRGC databases.

### 3.1. Databases

#### 3.1.1. AR database

The AR database [50] contains over 4000 color face images of 126 people (70 men and 56 women), including frontal views of faces with different facial expressions, lighting conditions and occlusions. In our experiments, 1400 frontal-face images captured across 2 sessions (separated by two weeks) of 100 subjects (50 males and 50 females) are selected. In each session, there are 7 undisguised images with different facial expressions and lighting conditions for each subject. The face region is cropped from the original images and resized to $56 \times 40$, which is similar to the image size used in [53]. Fig. 4 shows some cropped images from the AR database. We randomly select images of 10 subjects from 100 subjects in the first session to do color space training and images from the second session are used for testing.

#### 3.1.2. GT Database

The Georgia Tech (GT) [51] Face Database consists of 750 color images of 50 subjects (15 images per subject). These images have large variations in both pose and expression and some illumination changes. The face region is cropped from the original images and rescaled to the size of $30 \times 30$ similar to that used in [54]. The first eight images of randomly selected 10 from 50 subjects are used for color space training and the remaining seven images of all subjects serve as testing images. Fig. 5 shows some cropped images of the GT database.

#### 3.1.3. FRGC database

Face images in the FRGC database [34] are partitioned into three datasets, i.e., training, target and query datasets. There are 12,776 controlled or uncontrolled images in the training set, 16,028 controlled images in the target set, and 8014 uncontrolled images in the query set. The controlled images have good quality, while the uncontrolled images display poor quality, such as large illumination variations, low resolution, and blurring. These uncontrolled factors pose grand challenges to the face verification task.

**Fig. 6.** Cropped images of the FRGC database.

FRGC Experiment 4 has been reported to be the most challenging FRGC experiment [24], so it is chosen to assess the face recognition performance in our experiments. The face region is cropped from the original images and resized to a spatial resolution of $32 \times 32$, which is the same as that used in [17,23]. Following the FRGC protocol, we use the training set for color space training and the other two sets for testing. Fig. 6 shows some cropped images of the FRGC database.

### 3.2. The dependence of face recognition performance on the correlation between two chrominance components

Here, we conduct experiments to investigate the dependence of the face recognition performance on the correlation of two chrominance components $C_1$, $C_2$. With $\mathbf{u}_1$ determined in (27), we rotate $\mathbf{u}_2$ (28) in the 2-dimensional chrominance subspace so that different covariances (17) of $C_1$, $C_2$ and different $LuC_1C_2$ color spaces can be obtained, corresponding to different angles between $\mathbf{u}_1$ and $\mathbf{u}_2$.

We apply ERE [55] to raw pixels of $LuC_1C_2$ images and calculate the face recognition/verification rate. For AR database, the covariance of two chrominance components is plotted against $\delta$ (the angle between $\mathbf{u}_1$ and $\mathbf{u}_2$) on Fig. 7(a) and the face recognition rates achieved by different $LuC_1C_2$ color spaces are plotted against $\delta$ on Fig. 7(b). Similarly for the FRGC database, the covariance of $C_1$, $C_2$ is plotted against $\delta$ on Fig. 8(a) and face verification rates achieved by different $LuC_1C_2$ color spaces are plotted against $\delta$ on Fig. 8(b).

As we can observe on Figs. 7 and 8, the curve of covariance is very similar to that of the face recognition performance on both FRGC and AR databases. This observation provides evidence of the close dependence of the face recognition performance on the correlation of two chrominance components $C_1$, $C_2$. Moreover, the covariance will be maximized and the face recognition or verification rate will reach its minimum point when $\delta$ is around 0° ($\mathbf{u}_2$ is identical to $\mathbf{u}_1$) on both AR and FRGC datasets. The covariance will be minimized to 0 and the face recognition or verification rate will reach its maximum point when $\delta$ is around 100° on AR dataset and 60° on FRGC dataset rather than 90° ($\mathbf{u}_2$ is orthogonal to $\mathbf{u}_1$). Therefore, in order to achieve the best face recognition performance, we should minimize the correlation between $C_1$, $C_2$ rather than simply choose the $\mathbf{u}_2$ which is orthogonal to $\mathbf{u}_1$ as in [17].

### 3.3. Performance comparison of different color spaces under various conditions

In this experimental part, we follow the color face recognition framework shown on Fig. 9. A translated, rotated, cropped and resized color face image is firstly transformed from the RGB color space to another color space such as $LuC_1C_2$. Many recent color face recognition approaches conduct experiments on raw pixels [8,15,17,24]. Also, Gabor wavelet (GW) has been proven to be highly discriminative for FR [56]. Thus these two representative features are extracted from face images in the new color space.

In order to compare the face recognition performance of $LuC_1C_2$ with those of the other color spaces, three popular dimensionality reduction methods, i.e., PCA, enhanced fisher linear discriminant model (EFM) [57] and eigenfeature regularization and extraction (ERE) [55], are applied to the extracted color features. Also note

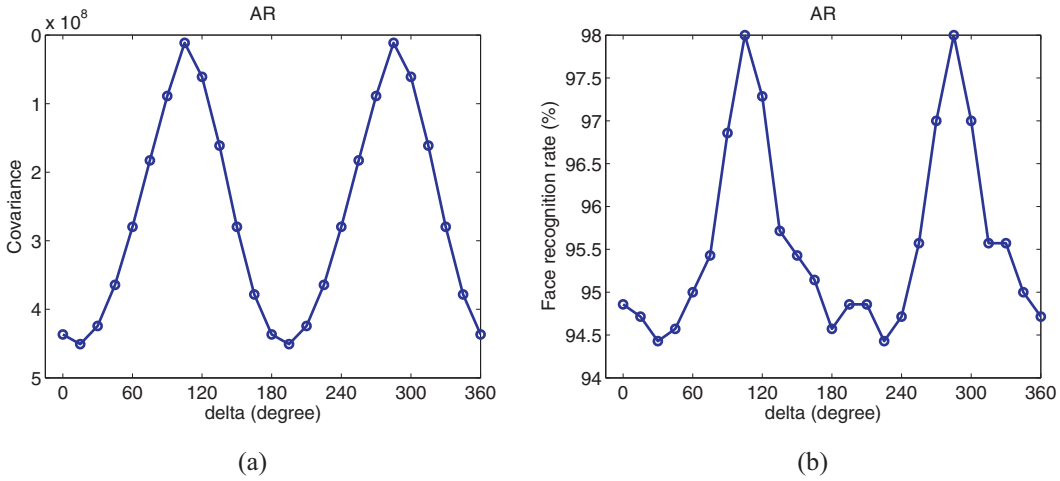



**Fig. 7.** (a) The covariance between two chrominance components plotted against the angle between $\mathbf{u}_1$, $\mathbf{u}_2$ and (b) corresponding face recognition rates on the AR database.

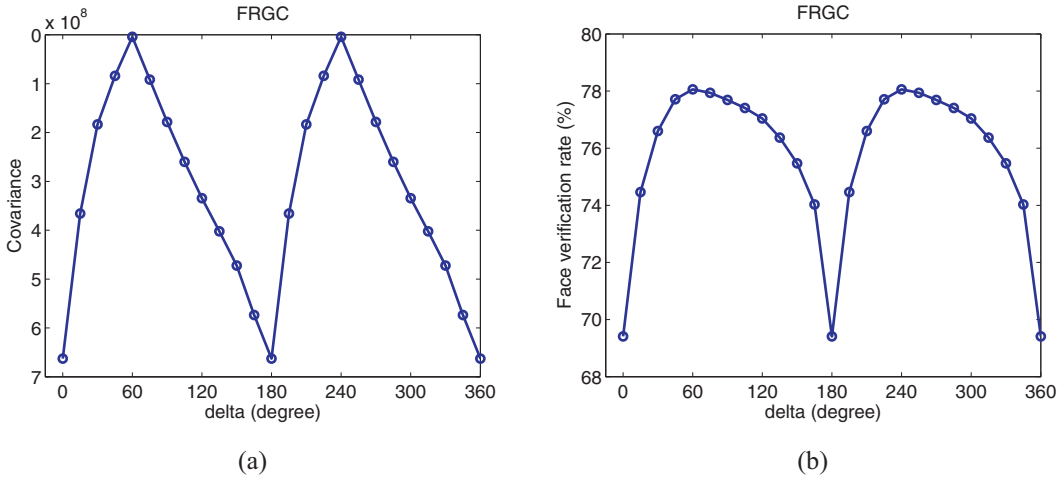



**Fig. 8.** (a) The covariance between two chrominance components plotted against the angle between $\mathbf{u}_1$, $\mathbf{u}_2$ and (b) corresponding face verification rates on the FRGC database.

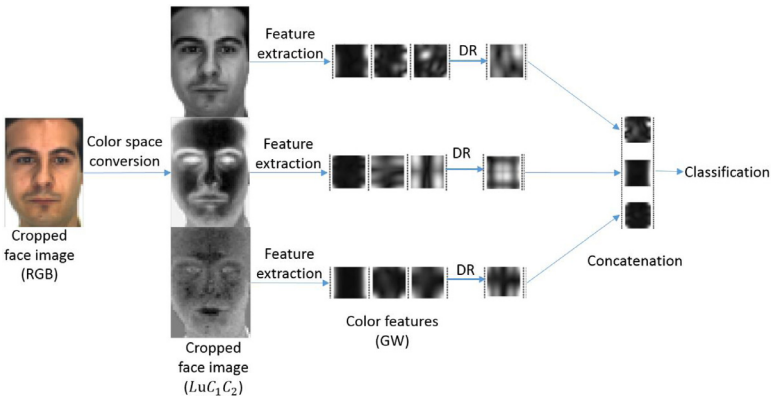



**Fig. 9.** The color face recognition framework used on AR and GT databases.

that PCA is commonly used as a benchmark for the evaluation of the performance in FR algorithms [22] and it may significantly enhance the recognition accuracy [58,59]. Plenty of color face recognition methods employ EFM method for low-dimensional feature extraction [15,17,23]. ERE outperforms all other FR methods discussed in [25,55,60].

According to [18,25,61,62], different color channels are processed separately and then concatenated together into a pattern vector for classification. So dimensionality reduction is implemented separately in 3 color channels on AR and GT. The low-dimensional color features from three color channels are concatenated into one vector to combine the information in three color channels. Note that all low-dimensional features would be normalized by removing the mean and dividing the standard deviation of the feature in each color channel before the final feature concatenation to avoid the negative effect of magnitude dominance of one component over the others [17]. After that, the nearest-neighbour classifier is used to classify all query images, where the Maha-

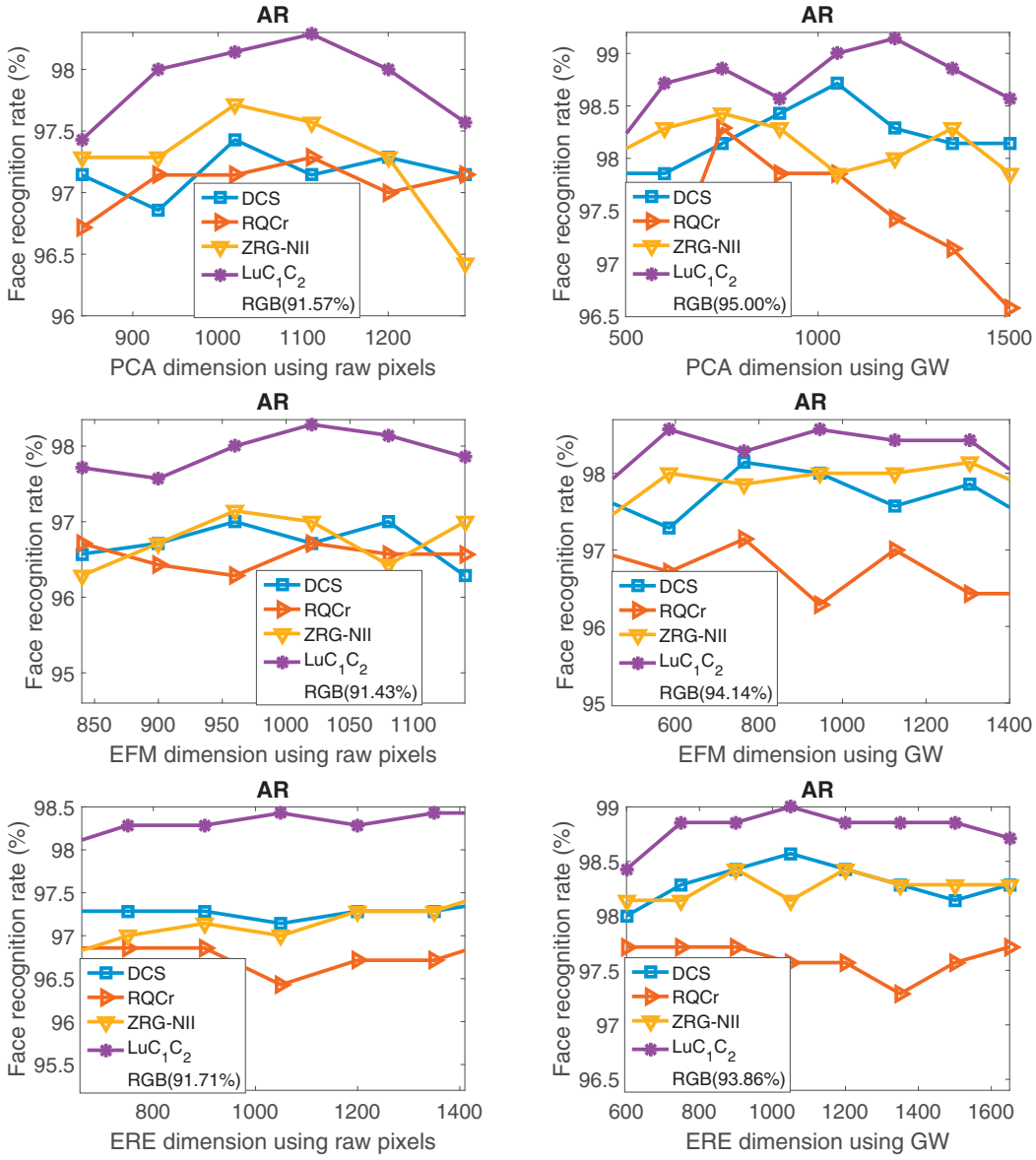



**Fig. 10.** Face recognition rates against feature dimension on AR, each column specifies one type of feature (2 in total) and each row specifies one dimension reduction method (3 in total).

lanobis distance is used for PCA and ERE, and the Cosine distance is used for EFM.

### 3.3.1. Results on AR

The face recognition rate (FRR), which is the ratio of the number of correctly classified query images to the total number of query images, is plotted against the feature dimension to present the overall face recognition performances of different color spaces on AR and GT databases. Note that, in our paper, the PCA dimension indicates the number of principle components of PCA, the EFM dimension indicates the number of principal components used in the PCA step of EFM, and the ERE dimension indicates the parameter m, which is the dimension of significant face components in [55]. As the face recognition rate of the RGB color space is far below those of the other color spaces, it is not plotted in the figures. Instead, we show the best FRR of the RGB color space among all feature dimensions in the legend area such as "RGB(FRR(%))". Using two color features and three dimension reduction methods, the face recognition rates of different color spaces are shown on Fig. 10. The results show that our proposed $LuC_1C_2$ color space performs better than three state-of-the-art color spaces consistently over all tested feature dimensions of two different features and three different dimension reduction methods.

### 3.3.2. Results on GT

Fig. 11 shows the face recognition rates of 5 color spaces using two color features and three dimension reduction methods on the GT database. Again, our proposed $LuC_1C_2$ color space achieves consistently better face recognition performance than state-of-the-art color spaces over all tested feature dimensions of two different features and three different dimension reduction methods.

### 3.3.3. Results on FRGC

FRGC is the most frequently-used database to evaluate the performances of different color spaces for face recognition. For direct comparisons with state-of-the-art results reported by other researchers, the commonly used evaluation framework on FRGC in

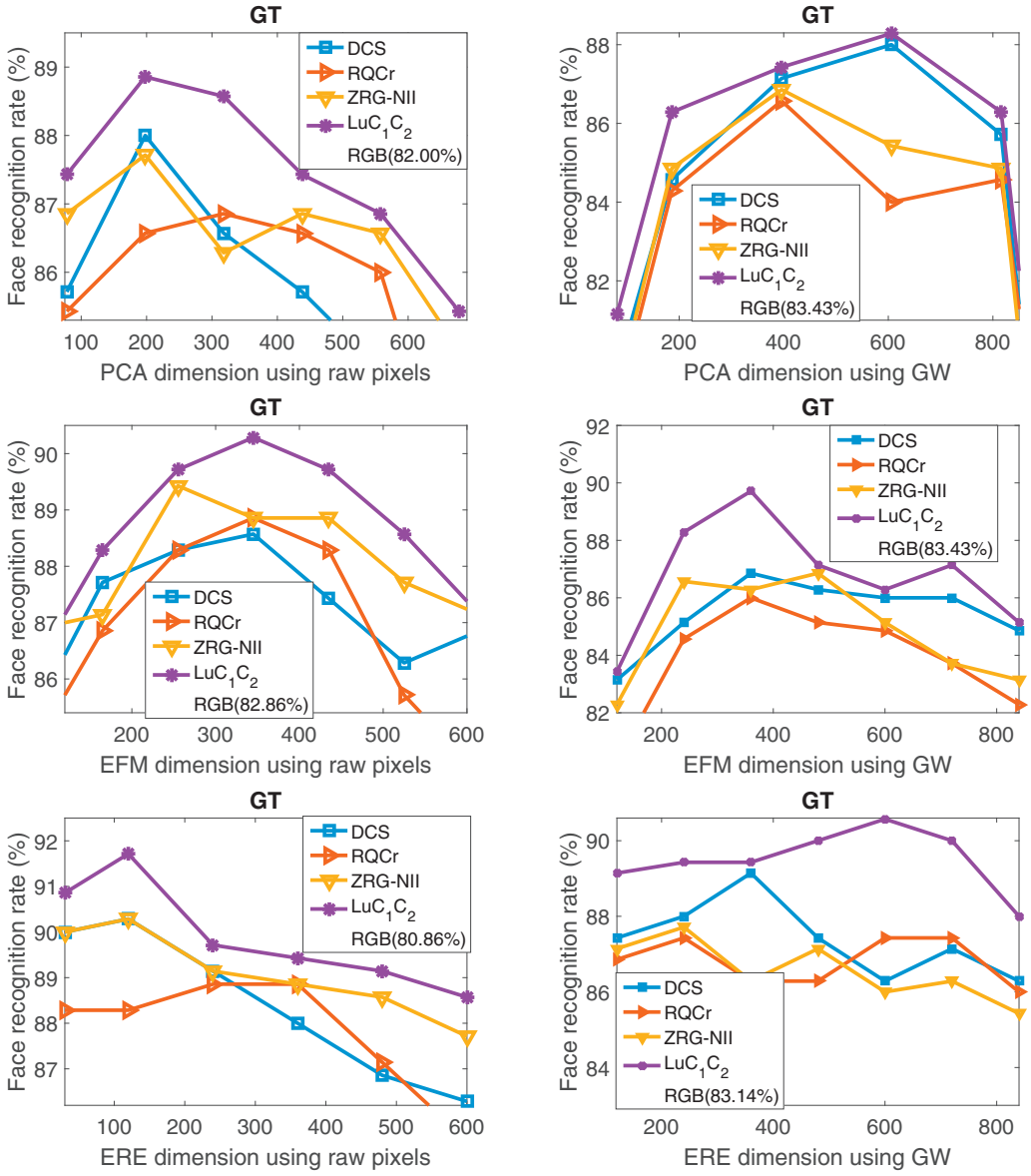



**Fig. 11.** Face recognition rates against feature dimension on GT, each column specifies one type of feature (2 in total) and each row specifies one dimension reduction method (3 in total).

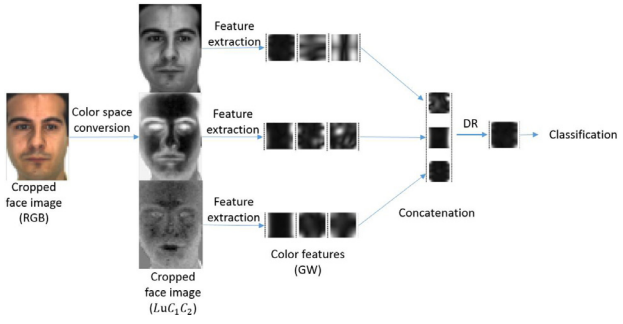



**Fig. 12.** The color face recognition framework used on FRGC databases.

[17,23] is adopted, as shown on Fig. 12. This guarantees fair and reliable comparisons of our proposed $LuC_1C_2$ color space with other color spaces.

The face recognition performance on FRGC is reported by plotting the face verification rate (FVR) at FAR=0.1% against the feature dimension so as to be consistent with recently published approaches [17,25]. The best FVR of the RGB color space among all feature dimensions is shown in the legend area such as "RGB(FVR(%))". As we can see from Fig. 13, for both of the two color features and the three dimension reduction methods, the face verification rates of the proposed $LuC_1C_2$ color space are consistently higher than those of DCS, ZRG-NII, RQCr and RGB color spaces over all tested feature dimensions. Note that the GW feature is extracted from face images of $64 \times 64$ rather than $32 \times 32$, which shows quite poor FR performance. The decision-level feature fusion method in [63] is used to fuse GW features extracted from 3 color channels.

### 3.4. Performance comparison of the proposed $LuC_1C_2$ color space with the state of the arts on FRGC

Many recent color face recognition approaches evaluate the effectiveness of different color spaces on raw pixels [8,15,17,24]. Table 1 shows results cited directly from published papers

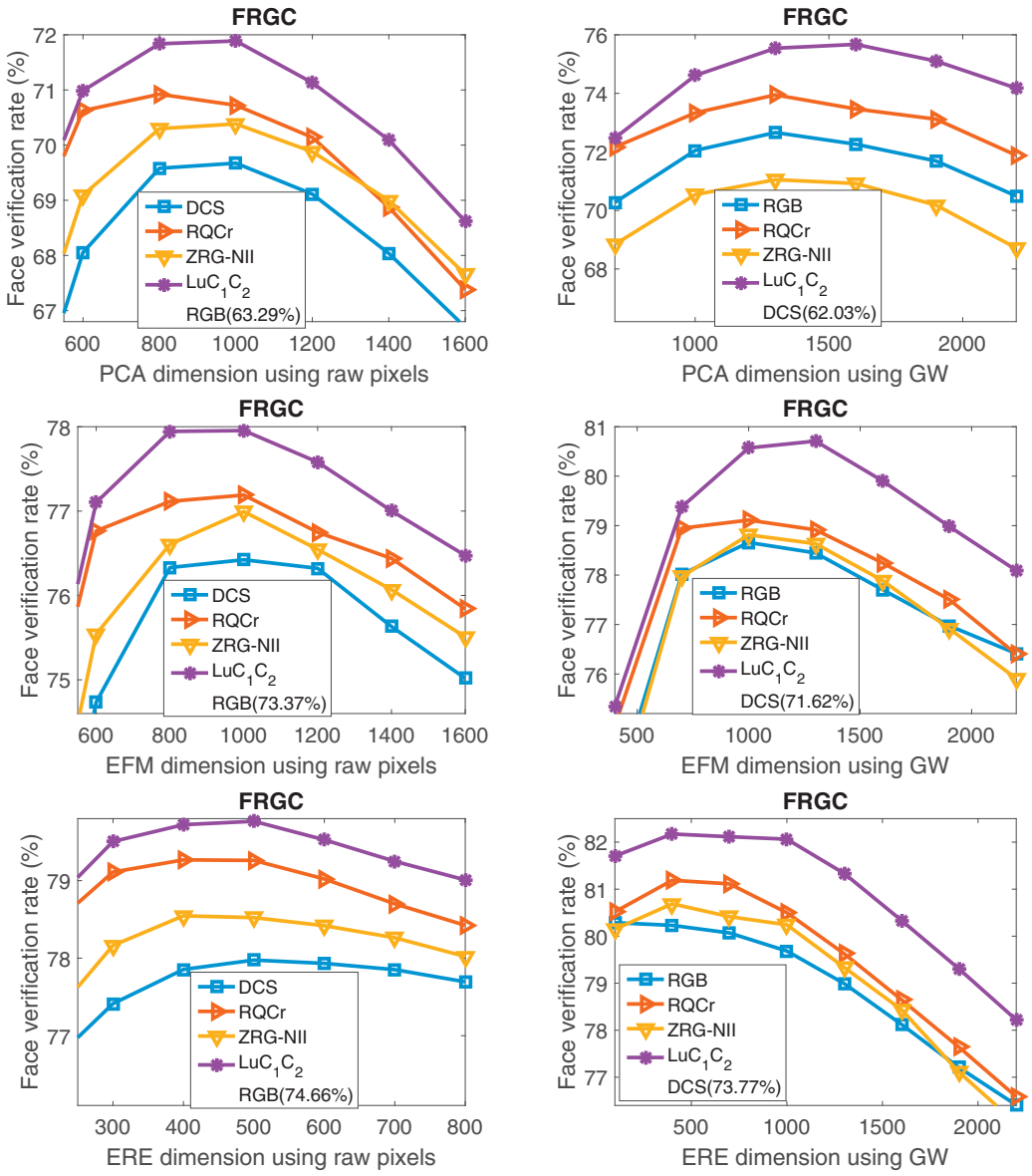



**Fig. 13.** Face verification rates against feature dimension on FRGC, each column specifies one type of feature (2 in total) and each row specifies one dimension reduction method (3 in total).

**Table 1**
Performance comparison of $LuC_1C_2$ with the state of the arts using raw pixels on FRGC.

| Color space | FVR (%) at FAR = 0.1% |
|---|---|
| Gray [17] | 36.94 |
| RGB [17] | 45.20 |
| YUV [17] | 67.29 |
| YIQ [17] | 69.75 |
| RQCr [17] | 72.51 |
| ZRG-NII [17] | 72.86 |
| DCS [23] | 76.26 |
| **$LuC_1C_2$** | **77.95** |

[17,23] on the FRGC database. Experimental settings are all set as below: raw pixels are used as features and EFM method is used for dimension reduction, where the feature dimension was set to 1000. Table 1 shows that (1) the use of color information greatly improves the FVR compared to the same scheme using the gray information only; (2) different color spaces possess significantly different classification capabilities for face recognition; (3) our proposed $LuC_1C_2$ color space performs significantly better than recently proposed color spaces on the FRGC database.

In recent years, many complex features and different feature fusion methods have been employed to achieve state-of-the-art results on FRGC. In [64], authors fuse local patterns of Gabor magnitude and phase using block-based Fisher's linear discriminant. Gabor and LBP features are processed by kernel LDA and then fused by decision-level fusion in [63]. In [65], multiscale local phase quantization and multiscale LBP are fused using kernel fusion. Multi-directional multi-level dual-cross patterns are fused by PLDA and score averaging in [66]. Authors in [67] use EFM to extract features from the real part, the imaginary part, and the magnitude of RIQ color images in the frequency domain. The features are then fused by means of concatenation. In [16], patch-based Gabor, multi-resolution LBP and component-based DCT features are extracted from R, Q, Cr images, respectively. EFM and the decision-level fusion approach are used to fuse different features. Authors in [68] combine multiple features extracted from both color norm

**Table 2**
Performance comparison of $LuC_1C_2$ with state of the arts using complex features on FRGC.

| Color space | FVR (%) at FAR = 0.1% |
|---|---|
| Gray [64] | 85.2 |
| Gray [63] | 88.1 |
| Gray [65] | 91.59 |
| Gray [66] | 92.89 |
| RIQ [67] | 80.03 |
| RQCr [16] | 92.43 |
| ZRG-NII [68] | 93.17 |
| **$LuC_1C_2$** | **93.38** |

patterns and color angular patterns of the ZRG-NII color space by feature concatenation. We can see that the above state-of-the-arts use multiple different features and different feature fusion schemes for different color spaces.

To compare the FVR of our color space $LuC_1C_2$ with various state of the arts, we extract raw-pixel, GW and LBP features from $LuC_1C_2$ images and simply fuse them at the decision level with equal weights. Table 2 tabulates the face verification performance of various approaches. All the state-of-the-art results are directly cited from recently published papers. Table 2 shows that the proposed $LuC_1C_2$ color space achieves higher FVR than all reported results on FRGC database.

### 3.5. The combination of CNN features with those from the $LuC_1C_2$ color space

Recently, some deep learning approaches are developed for face recognition based on convolutional neural networks (CNN) [69]. Although high face recognition accuracy can be achieved by CNN, it needs a huge number (millions) of training images and a very complex training process with great efforts in selecting/adjusting training parameters. Sophisticatedly trained CNN achieve high accuracy for some testing datasets but may not perform well for other specific applications. The common solution is to fine-tune the pre-trained models by application-specific training data [70–72]. Motivated by DeepID in [73], which learns features from both RGB images and gray-scale images, we enhance the CNN performance by using simple raw pixels in the proposed color space. The VGG-Face model in [74] is designed for face recognition [75,76] and pre-trained by 2.6 million RGB face images. The VGG-Face model has been widely used by researchers to extract CNN features from face images [70–72]. Thus we adopt this model for CNN feature extraction from RGB images in the experiments. Note that the 4096-dimensional feature vector at layer 'fc6' of the VGG-Face model is taken as the CNN feature. We calculate the similarity between two CNN features using cosine distance metric as in [77,78]. The VGG-Face model achieves the face verification accuracy of 97.27% on LFW [52] under this setting. To enhance the performance of CNN, we simply apply PCA to the raw-pixel feature in the proposed color space and fuse it with CNN features at the decision level.

#### 3.5.1. Results on LFW

The LFW database [52] has been widely used as a benchmark dataset to evaluate face recognition algorithms using deep neural networks [73]. It consists of 13,233 images of 5749 subjects. All images are collected from the internet with large pose, illumination and facial-expression variations, as well as occlusions. Fig. 14 shows some cropped images of the LFW database.

We report the averaged results over 10 folds of View 2 as in [73,74,79]. More precisely, 3000 positive pairs (two images of the same subject) and 3000 negative pairs (two images of different subjects) are divided into 10 subsets. In each fold, 9 subsets are used for training and the hold-out subset is used for testing. The task is to determine whether an image pair comes from the same subject or not. Both images with and without backgrounds are used to test the face verification performance of different approaches. As there is no label information to use for fine-tuning the VGG-Face model on LFW, only the pre-trained VGG-Face model is tested. Results are shown on Table 3. As we can see on Table 3, CNN+$LuC_1C_2$ increases the face verification accuracy of CNN by 1.7% for images without backgrounds. Simple raw pixels of $LuC_1C_2$ only marginally improve the face verification accuracy of CNN for images with backgrounds, because the VGG-Face model was initially trained and tuned to achieve very good results for this type of images.

**Fig. 14.** Cropped images of the LFW database.

**Table 3**
Face verification accuracy (%) using CNN and raw pixels of $LuC_1C_2$ color space on LFW.

| Background | Model | CNN | CNN+$LuC_1C_2$ |
|---|---|---|---|
| Without | VGG-Face | 85.62 | **87.32** |
| With | VGG-Face | 97.43 | **97.72** |

**Table 4**
Face verification rate (%) at FAR = 0.1% using CNN and raw pixels of $LuC_1C_2$ color space on FRGC.

| Background | Model | CNN | CNN+$LuC_1C_2$ |
|---|---|---|---|
| Without | VGG-Face | 45.32 | **89.07** |
| | FVGG-Face | 75.34 | **93.42** |
| With | VGG-Face | 88.86 | **96.98** |
| | FVGG-Face | 92.50 | **97.57** |

#### 3.5.2. Results on FRGC

Both images with and without backgrounds are used in the experiments. To improve the FR performance of the VGG-Face model on FRGC database, we fine-tune the VGG-Face model using images from the FRGC training set. As there are 222 subjects in the training set, we change the output of the last fully connected layer from 2622 to 222. After fine-tuning, we obtain the fine-tuned VGG-Face model, FVGG-Face model. In the experiments, both the pre-trained VGG-Face model and the fine-tuned FVGG-Face model are used. Results are shown on Table 4.

Directly applying the pre-trained VGG-Face model to the FRGC database provides unsatisfactory results as shown on Table 4. The widely used fine-tuning can only increase the FVR of CNN to 75.34% and 92.50%. The effect of fine-tuning on pre-trained VGG-Face is limited because fine-tuning is still based on the pre-trained model. Pixel values in our proposed $LuC_1C_2$ color space are complementary to the CNN features extracted from RGB images. Therefore, CNN+$LuC_1C_2$ greatly improves the FVR of VGG-Face model to 89.07% and 96.98% for images without and with backgrounds, respectively. These performance improvements are even more significant than those achieved by fine-tuning. Furthermore, CNN+$LuC_1C_2$ improves the accuracy of the fine-tuned VGG-Face model, FVGG-Face model, to 93.42% and 97.57% for images without and with backgrounds, respectively.

#### 3.5.3. Discussion

Although the complex deep CNN models might contain some built-in function of color space conversion, experimental results on Table 3 and Table 4 indicate that features from our proposed

**Table 5**
Face verification accuracy/rate (%) using traditional features from RGB and $LuC_1C_2$ color spaces on LFW and FRGC. RP indicates raw pixels and GW indicates Gabor wavelet.

| Dataset | RP-RGB | RP-RGB & RP-$LuC_1C_2$ | GW-RGB | GW-RGB & GW-$LuC_1C_2$ |
|---|---|---|---|---|
| LFW | 65.25 | **71.38** | 69.05 | **72.57** |
| FRGC | 71.85 | **80.42** | 72.66 | **89.74** |

$LuC_1C_2$ color space can be used to enhance the face verification performance of CNN on LFW and FRGC databases.

### 3.6. The combination of RGB features with those from the $LuC_1C_2$ color space

The proposed $LuC_1C_2$ color space can also be combined with the RGB color space to enhance the performance of traditional features. We conduct experiments on LFW and FRGC datasets. Traditional features of raw pixels or Gabor wavelet are extracted from RGB and $LuC_1C_2$ face images. Features from two different color spaces are then fused at the decision level. We report the face verification performance of different approaches on LFW and FRGC datasets on Table 5. The experimental results show that, the face verification performance can be significantly improved by combining the traditional features from the RGB and the proposed $LuC_1C_2$ color spaces.

## 4. Conclusion

In this paper, we propose a color space $LuC_1C_2$ based on a framework of constructing effective color spaces in providing high face recognition performance. It consists of one luminance component $Lu$ and two chrominance components $C_1$, $C_2$. The luminance component $Lu$ is chosen among 4 different luminance candidates by studying their R,G,B coefficients and the color sensor properties. For two chrominance components $C_1$, $C_2$, the directions of their transform vectors are computed by the discriminant analysis and the covariance analysis in the chrominance subspace of the RGB color space. The magnitudes of their transform vectors are computed according to the discriminant values of $Lu$, $C_1$, $C_2$. Extensive experiments are conducted on 4 benchmark databases. Experimental results using 2 distinct features and 3 different dimension reduction methods on 3 databases show that our proposed color space $LuC_1C_2$ achieves consistently better face recognition performance than state-of-the-art color spaces. The experiments also show that features extracted from the proposed color space achieve higher face verification rate than all published results on FRGC database. Furthermore, we show that features from the proposed color space can be used to improve the face verification performance of pre-trained CNN models, fine-tuned CNN models and traditional features.

## Acknowledgment

This research was carried out at the Rapid-Rich Object Search (ROSE) Lab at the Nanyang Technological University, Singapore. This research was partly supported by Singapore Ministry of Education Academic Research Fund Tier 1 RG 123/15. The ROSE Lab is supported by the Infocomm Media Development Authority, Singapore.

## References

[1] S. Bianco, G. Ciocca, C. Cusano, R. Schettini, Automatic color constancy algorithm selection and combination, Pattern Recognit. 43 (3) (2010) 695–705.
[2] S.W. Oh, S.J. Kim, Approaching the computational color constancy as a classification problem through deep learning, Pattern Recognit. 61 (2017) 405–416.
[3] G.H. Liu, J.Y. Yang, Content-based image retrieval using color difference histogram, Pattern Recognit. 46 (1) (2013) 188–198.
[4] M. Bressan, D. Guillamet, J. Vitria, Using an ICA representation of local color histograms for object recognition, Pattern Recognit. 36 (3) (2003) 691–701.
[5] R. Lukac, K.N. Plataniotis, Color Image Processing: Methods and Applications, CRC Press, 2006.
[6] C. Garcia, G. Tziritas, Face detection using quantized skin color regions merging and wavelet packet analysis, IEEE Trans. Multimed. 1 (3) (1999) 264–277.
[7] K. Cho, J. Jang, K. Hong, Adaptive skin-color filter, Pattern Recognit. 34 (5) (2001) 1067–1073.
[8] P. Shih, C. Liu, Comparative assessment of content-based face image retrieval in different color spaces, Int. J. Pattern Recognit. Artif. Intell. 19 (07) (2005) 873–893.
[9] L. Qiao, S. Chen, X. Tan, Sparsity preserving projections with applications to face recognition, Pattern Recognit. 43 (1) (2010) 331–341.
[10] A.A. Mohammed, R. Minhas, Q.J. Wu, M.A. Sid-Ahmed, Human face recognition based on multidimensional PCA and extreme learning machine, Pattern Recognit. 44 (10–11) (2011) 2588–2597.
[11] J. Gui, Z. Sun, W. Jia, R. Hu, Y. Lei, S. Ji, Discriminant sparse neighborhood preserving embedding for face recognition, Pattern Recognit. 45 (8) (2012) 2884–2893.
[12] O. Arandjelović, Colour invariants under a non-linear photometric camera model and their application to face recognition from video, Pattern Recognit. 45 (7) (2012) 2499–2509.
[13] L. Torres, J. Reutter, L. Lorente, The importance of the color information in face recognition, in: Proceedings of the IEEE International Conference on Image Processing, 3, 1999, pp. 627–631.
[14] M. Rajapakse, J. Tan, J. Rajapakse, Color channel encoding with NMF for face recognition, in: Proceedings of the IEEE International Conference on Image Processing, 3, 2004, pp. 2007–2010.
[15] J.Y. Choi, Y.M. Ro, K.N. Plataniotis, Color face recognition for degraded face images, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 39 (5) (2009) 1217–1230.
[16] Z. Liu, C. Liu, Fusion of color, local spatial and global frequency information for face recognition, Pattern Recognit. 43 (8) (2010) 2882–2890.
[17] J. Yang, C. Liu, L. Zhang, Color space normalization: enhancing the discriminating power of color spaces for face recognition, Pattern Recognit. 43 (4) (2010) 1454–1466.
[18] Z. Lu, X. Jiang, A.C. Kot, A color channel fusion approach for face recognition, IEEE Signal Process Lett. 22 (11) (2015) 1839–1843.
[19] J.Y. Choi, Y.M. Ro, K.N. Plataniotis, A comparative study of preprocessing mismatch effects in color image based face recognition, Pattern Recognit. 44 (2) (2011) 412–430.
[20] P. Shih, C. Liu, Improving the face recognition grand challenge baseline performance using color configurations across color spaces, in: Proceedings of the IEEE International Conference on Image Processing, 2006, pp. 1001–1004.
[21] M.T. Sadeghi, S. Khoshrou, J. Kittler, Confidence based gating of colour features for face authentication, in: Multiple Classifier Systems, Springer, 2007, pp. 121–130.
[22] P.J. Phillips, H. Moon, S. Rizvi, P.J. Rauss, et al., The FERET evaluation methodology for face-recognition algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 22 (10) (2000) 1090–1104.
[23] J. Yang, C. Liu, A discriminant color space method for face representation and verification on a large-scale database, in: Proceedings of the IEEE International Conference on Pattern Recognition, 2008, pp. 1–4.
[24] C. Liu, Learning the uncorrelated, independent, and discriminating color spaces for face recognition, IEEE Trans. Inf. Forensics Secur. 3 (2) (2008) 213–222.
[25] J.Y. Choi, Y.M. Ro, K.N. Plataniotis, Color local texture features for color face recognition, IEEE Trans. Image Process. 21 (3) (2012) 1366–1380.
[26] V. Georges, System of television in colors, 1945, US Patent 2375966.
[27] R.-L. Hsu, M. Abdel-Mottaleb, A.K. Jain, Face detection in color images, IEEE Trans. Pattern Anal. Mach. Intell. 24 (5) (2002) 696–706.
[28] J. Kittler, M.T. Sadeghi, Physics-based decorrelation of image data for decision level fusion in face verification, in: Multiple Classifier Systems, Springer, 2004, pp. 354–363.
[29] W.H. Buchsbaum, Color TV Servicing, Prentice Hall, 1975.
[30] P. Colantoni, et al., Color space transformations, see http://www.radugaryazan.ru/files/doc/colorspacetransform95.pdf (2004).
[31] A. Yip, P. Sinha, Role of color in face recognition(2001).
[32] J. Yang, C. Liu, A general discriminant model for color face recognition, in: Proceedings of the IEEE International Conference on Computer Vision, 2007, pp. 1–6.
[33] Y. Ohta, Knowledge-Based Interpretation of Outdoor Natural Color Scenes, 4, Morgan Kaufmann, 1985.

[34] 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, 1, 2005, pp. 947–954.
[35] J. Lu, K.N. Plataniotis, On conversion from color to gray-scale images for face detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2009, pp. 114–119.
[36] J. Yang, D. Zhang, A.F. Frangi, J.-y. Yang, Two-dimensional PCA: a new approach to appearance-based face representation and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 26 (1) (2004) 131–137.
[37] T. Ahonen, A. Hadid, M. Pietikainen, Face description with local binary patterns: application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 28 (12) (2006) 2037–2041.
[38] Y. Su, S. Shan, X. Chen, W. Gao, Hierarchical ensemble of global and local classifiers for face recognition, IEEE Trans. Image Process. 18 (8) (2009) 1885–1896.
[39] Wikipedia, Color vision — wikipedia, the free encyclopedia, 2016, [Online; Accessed 25.03.16].
[40] LibRaw, Channel noise and raw converters, 2008,
[41] R. Kawakami, Z. Hongxun, R.T. Tan, K. Ikeuchi, Camera spectral sensitivity and white balance estimation from sky images, Int. J. Comput. Vis. (2013).
[42] D. Wu, J. Tian, B. Li, Y. Wang, Y. Tang, Recovering sensor spectral sensitivity from raw data, J Electron. Imaging 22 (2) (2013) 023032.
[43] G.D. Finlayson, S.D. Hordley, P.M. Hubel, Color by correlation: a simple, unifying framework for color constancy, IEEE Trans. Pattern Anal. Mach. Intell. 23 (11) (2001) 1209–1221.
[44] X. Jing, Q. Liu, C. Lan, J. Man, S. Li, D. Zhang, Holistic orthogonal analysis of discriminant transforms for color face recognition, in: Proceedings of the IEEE Conference on Image Processing, IEEE, 2010, pp. 3841–3844.
[45] C. Zhao, D. Miao, Z. Lai, C. Gao, C. Liu, J. Yang, Two-dimensional color uncorrelated discriminant analysis for face recognition, Neurocomputing 113 (2013) 251–261.
[46] P.L. Duren, Theory of Hp Spaces, 38, IMA, 1970.
[47] H. Stokman, T. Gevers, Selection and fusion of color models for image feature detection, IEEE Trans. Pattern Anal. Mach. Intell. 29 (3) (2007) 371–381.
[48] B.V. Funt, G.D. Finlayson, Color constant color indexing, IEEE Trans. Pattern Anal. Mach. Intell. 17 (5) (1995) 522–529.
[49] G. Snedegor, W.G. Cochran, et al., Statistical methods, Statistical Methods. (6th ed), The Iowa State University Press, 1967.
[50] A.M. Martinez, The AR face database, CVC Technical Report, CVC 24(1998).
[51] A. Nefian, A. Nefian, Georgia tech face database, 2013,
[52] G.B. Huang, M. Ramesh, T. Berg, E. Learned-Miller, Labeled faces in the wild: a database for studying face recognition in unconstrained environments, Technical report 07–49, University of Massachusetts, Amherst, 2007.
[53] X. Jiang, J. Lai, Sparse and dense hybrid representation via dictionary decomposition for face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 37 (5) (2015) 1067–1079.
[54] S.-J. Wang, J. Yang, N. Zhang, C.-G. Zhou, Tensor discriminant color space for face recognition, IEEE Trans. Image Process. 20 (9) (2011) 2490–2501.
[55] X. Jiang, B. Mandal, A. Kot, Eigenfeature regularization and extraction in face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 30 (3) (2008) 383–394.
[56] C. Liu, H. Wechsler, Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition, IEEE Trans. Image Process. 11 (4) (2002) 467–476.
[57] C. Liu, H. Wechsler, Robust coding schemes for indexing and retrieval from large face databases, IEEE Trans. Image Process. 9 (1) (2000) 132–137.
[58] X. Jiang, Linear subspace learning-based dimensionality reduction, IEEE Signal Process Mag. 28 (2) (2011) 16–26.
[59] X. Jiang, Asymmetric principal component and discriminant analyses for pattern classification, IEEE Trans. Pattern Anal. Mach. Intell. 31 (5) (2009) 931–937.
[60] J.Y. Choi, W. De Neve, K.N. Plataniotis, Y.M. Ro, Collaborative face recognition for improved face annotation in personal photo collections shared on online social networks, IEEE Trans. Multimed. 13 (1) (2011) 14–28.
[61] C. Liu, Extracting discriminative color features for face recognition, Pattern Recognit Lett. 32 (14) (2011) 1796–1804.
[62] J.Y. Choi, Y.M. Ro, K.N. Plataniotis, Boosting color feature selection for color face recognition, IEEE Trans. Image Process. 20 (5) (2011) 1425–1434.
[63] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE Trans. Image Process. 19 (6) (2010) 1635–1650.
[64] S. Xie, S. Shan, X. Chen, J. Chen, Fusing local patterns of Gabor magnitude and phase for face recognition, IEEE Trans. Image Process. 19 (5) (2010) 1349–1361.
[65] C.H. Chan, M.A. Tahir, J. Kittler, M. Pietikäinen, Multiscale local phase quantization for robust component-based face recognition using kernel fusion of multiple descriptors, IEEE Trans. Pattern Anal. Mach. Intell. 35 (5) (2013) 1164–1177.
[66] C. Ding, J. Choi, D. Tao, L.S. Davis, Multi-directional multi-level dual-cross patterns for robust face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 38 (3) (2016) 518–531.
[67] Z. Liu, C. Liu, A hybrid color and frequency features method for face recognition, IEEE Trans. Image Process. 17 (10) (2008) 1975–1980.
[68] S.H. Lee, J.Y. Choi, Y.M. Ro, K.N. Plataniotis, Local color vector binary patterns from multichannel face images for face recognition, IEEE Trans. Image Process. 21 (4) (2012) 2347–2353.
[69] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Proceedings of the Advances in Neural Information Processing systems, 2012, pp. 1097–1105.
[70] F. Gurpinar, H. Kaya, H. Dibeklioglu, A. Salah, Kernel elm and CNN based facial age estimation, in: Proceedings of the The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2016.
[71] K. Zhang, L. Tan, Z. Li, Y. Qiao, Gender and smile classification using deep convolutional neural networks, in: Proceedings of the The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2016.
[72] X. Peng, Z. Xia, L. Li, X. Feng, Towards facial expression recognition in the wild: a new database and deep recognition system, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016, pp. 93–99.
[73] Y. Sun, X.G. Wang, X.O. Tang, Deep learning face representation from predicting 10,000 classes, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1891–1898.
[74] O.M. Parkhi, A. Vedaldi, A. Zisserman, Deep face recognition, in: Proceedings of the British Machine Vision Conference, 1, 2015, p. 6.
[75] http://www.robots.ox.ac.uk/~vgg/software/vggface/, 2015,
[76] X. Wu, R. He, Z. Sun, A lightened CNN for deep face representation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015.
[77] J.C. Chen, V.M. Patel, R. Chellappa, Unconstrained face verification using deep CNN features, in: Proceedings of the IEEE Winter Conference on Applications of Computer Vision, 2016, pp. 1–9.
[78] C. Ding, D. Tao, Robust face recognition via multimodal deep face representation, IEEE Trans. Multimed. 17 (11) (2015) 2049–2058.
[79] D. Chen, X. Cao, F. Wen, J. Sun, Blessing of dimensionality: high-dimensional feature and its efficient compression for face verification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 3025–3032.

**Ze Lu** received the B.Eng. degree from the Zhejiang University, China, in 2013. He is currently doing his Ph.D. degree in Nanyang Technological University, Singapore. His research interests include face recognition, dimension reduction and computer vision.

**Xudong Jiang** received the B.Eng. and M.Eng. degrees from the University of Electronic Science and Technology of China (UESTC) in 1983 and 1986, respectively, and the Ph.D. degree from the University of German Federal Armed Forces, Hamburg, Germany, in 1997, all in electrical and electronic engineering. From 1986 to 1993, he was a lecturer at UESTC, where he received two Science and Technology Awards from the Ministry for Electronic Industry of China. From 1993 to 1997, he was with the University of German Federal Armed Forces, as a scientific assistant. From 1998 to 2002, he was with Nanyang Technological University (NTU), Singapore, as a senior research fellow, where he developed a fingerprint verification algorithm that achieved the most efficient and the second most accurate fingerprint verification at the International Fingerprint Verification Competition (FVC 00). From 2002 to 2004, he was a lead scientist and the head of the Biometrics Laboratory at the Institute for Infocomm Research, Singapore. He joined NTU as a faculty member in 2004. Currently, he serves as the director of the Centre for Information Security, the School of Electrical and Electronic Engineering, NTU, Singapore. His research interest includes pattern recognition, signal and image processing, computer vision, and biometrics. He is a senior member of the IEEE.

**Alex Kot** received the BSEE and MBA degrees from the University of Rochester, New York, and the Ph.D. degree from the University of Rhode Island, Kingston. He was with AT&T briefly. Since 1991, he has been with Nanyang Technological University, Singapore. He headed the Division of Information Engineering at the School of Electrical and Electronic Engineering for eight years and is currently a professor and the vice dean (Research) of the School of Electrical and Electronic Engineering. He has published extensively in the areas of signal processing for communication, biometrics, data hiding, and authentication. He has served as an associate editor for IEEE Transactions on Signal Processing, IEEE Transactions on Circuits and Systems II, and IEEE Transactions on Circuits and Systems for Video Technology. He has served as a guest editor for IEEE and EURASIP journals. Currently, he is an associate editor for the IEEE Transactions on Circuits and Systems I and EURASIP Journal on Applied Signal Processing. He has served on numerous conference committees and technical committees, including co chairing the IEEE International Conference on Image Processing (ICIP) in 2004 and an IEEE distinguished lecturer (2005–2006). He is a fellow of the IEEE.