International Journal of Pattern Recognition
and Artificial Intelligence
Vol. 19, No. 7 (2005) 873–893

World Scientific
www.worldscientific.com

COMPARATIVE ASSESSMENT OF CONTENT-BASED FACE IMAGE RETRIEVAL IN DIFFERENT COLOR SPACES

PEICHUNG SHIH* and CHENGJUN LIU†

Department of Computer Science
New Jersey Institute of Technology
Newark, NJ 07102, USA
**ps9@oak.njit.edu*
†*liu@cs.njit.edu*

Content-based face image retrieval is concerned with computer retrieval of face images (of a given subject) based on the geometric or statistical features automatically derived from these images. It is well known that color spaces provide powerful information for image indexing and retrieval by means of color invariants, color histogram, color texture, etc. This paper assesses comparatively the performance of content-based face image retrieval in different color spaces using a standard algorithm, the Principal Component Analysis (PCA), which has become a popular algorithm in the face recognition community. In particular, we comparatively assess 12 color spaces (RGB, HSV, YUV, $YCbCr$, XYZ, YIQ, $L^*a^*b^*$, $U^*V^*W^*$, $L^*u^*v^*$, $I_1I_2I_3$, HSI, and rgb) by evaluating seven color configurations for every single color space. A color configuration is defined by an individual or a combination of color component images. Take the RGB color space as an example, possible color configurations are R, G, B, RG, RB, GB and RGB. Experimental results using 600 FERET color images corresponding to 200 subjects and 456 FRGC (Face Recognition Grand Challenge) color images of 152 subjects show that some color configurations, such as YV in the YUV color space and YI in the YIQ color space, help improve face retrieval performance.

Keywords: Color space; content-based face image retrieval; Principal Component Analysis (PCA); RGB; HSV; YUV; $YCbCr$; XYZ; YIQ; $L^*a^*b^*$; $U^*V^*W^*$; $L^*u^*v^*$; $I_1I_2I_3$; HSI; rgb.

1. Introduction

Content-based face image retrieval is concerned with computer retrieval of face images (of a given subject) based on the geometric or statistical features automatically derived from these images.[4,24,34] Efficient retrieval requires a robust feature extraction method that has the ability to learn meaningful low-dimensional patterns in spaces of very high dimensionality. Low-dimensional representations are also important when one considers the intrinsic computational aspect. The Principal Component Analysis (PCA)[6] has been widely used to perform dimensionality reduction for face indexing and retrieval.[14,15,32,33] In particular, PCA is the method

behind the Eigenfaces coding scheme[28] whose primary goal is to project the similarity judgment for face recognition onto a low-dimensional space. This space defines a feature space, or a "face space", which drastically reduces the dimensionality of the original space, and face detection and identification are carried out in this reduced face space.

It is well known that color spaces provide powerful information for image indexing and retrieval by means of color invariants, color histogram, color texture, etc. We assess, in this paper, the performance of content-based face image retrieval in different color spaces using a standard algorithm, PCA.[6] Specifically, we assess comparatively 12 color spaces (RGB, HSV, YUV, $YCbCr$, XYZ, YIQ, $L^*a^*b^*$, $U^*V^*W^*$, $L^*u^*v^*$, $I_1I_2I_3$, HSI, and rgb) by evaluating seven color configurations for every single color space. A color configuration is defined by an individual or a combination of color component images. Take the RGB color space as an example, possible color configurations are R, G, B, RG, RB, GB and RGB. Note that when two or three color component images are used to define a color configuration, each color component image is first normalized to zero mean and unit variance, and then the normalized color component images are concatenated to form an augmented vector representing the color configuration. Experimental results using 600 FERET color images corresponding to 200 subjects and 456 FRGC (Face Recognition Grand Challenge) color images of 152 subjects show that some color configurations, such as YV in the YUV color space and YI in the YIQ color space, help improve face retrieval performance.

2. Background

Color provides powerful information for object detection, indexing and retrieval.[12,31] Color histograms and color invariant moments provide robust object recognition against image variations such as illumination.[5,10] Swain and Ballard[21] developed a color indexing system which applies color histogram for image inquiry from a large image database.[25] The system separates the chrominance from the luminance and the color information derived is invariant to illumination variations.

In general, different color spaces, which are defined by means of transformations from the original RGB (red, green, blue) color space, display different color properties. The HSV (hue, saturation, value) color space and its variants, such as the HSI (hue, saturation, intensity) color space and the HLS (hue, lightness, saturation) color space, are often applied in locating and extracting facial features.[7,23] The $YCbCr$ (luminance, Chrominance-blue, Chrominance-red) color space, the YIQ (luminance, in-phase, quadrature) color space, and the YUV color space have wide applications in color clustering and quantization for skin color regions.[7,9,20,30] The perceptually uniform color spaces, such as the CIE-$U^*V^*W^*$ color space, the CIE-$L^*u^*v^*$ color space, and the CIE-$L^*a^*b^*$ color space have general and ubiquitous applications.[16,26,29,35]

The color information provided by the different color spaces, thus, can be applied for different visual tasks. Skin color, for example, has been used for face detection.[11,12,26] Yang *et al.*[31] suggested that the luminance account for most of the skin color variation and showed that the color histogram based on normalized red and green color components occupies a small cluster in the histogram. Sobottka *et al.*,[23] on the other hand, showed that the HSV color space and its variations have the advantage of providing large variance among facial features and are suitable for locating and extracting facial features. Hsu *et al.*[12] applied a nonlinear color transformation to detect skin patches in order to find human faces in an image. Torres *et al.*,[27] however, showed that color information does not improve the face recognition performance in comparison with the intensity information.

3. Color Spaces

This section details the 12 color spaces assessed in this paper: the RGB color space, the rgb color space, the $I_1I_2I_3$ (decorrelated RGB) color space, the CIE-XYZ color space, the human perceptual color spaces HSV and HSI, the video transmission efficiency color spaces YIQ, YUV and $YCbCr$, and the CIE perceptually uniform color spaces CIE-$U^*V^*W^*$, CIE-$L^*u^*v^*$, and CIE-$L^*a^*b^*$.

3.1. *The rgb color space*

The RGB images are sensitive to luminance, surface orientation, and other photographic conditions. To minimize such sensitivity, one can project the R, G, B values onto the $R = G = B = \max\{R, G, B\}$ plane. The projection spans a normalized rgb chromaticity triangle. The transformation is defined as follows:

$$\begin{aligned} r &= R/(R+G+B) \\ g &= G/(R+G+B) \\ b &= B/(R+G+B). \end{aligned} \tag{1}$$

3.2. *The $I_1I_2I_3$ color space*

Another approach to stabilize RGB images is to decorrelate the RGB components. The $I_1I_2I_3$ color space proposed by Ohta *et al.*[18] applies a Karhunen–Loeve transformation to decorrelate the RGB components. The linear transformation based on Ohta's experimental model is defined as follows[18]:

$$\begin{aligned} I_1 &= (R+G+B)/3 \\ I_2 &= (R-B)/2 \\ I_3 &= (2G-R-B)/2. \end{aligned} \tag{2}$$

3.3. *Human perceptual color spaces*

The HSV and the HSI color spaces are motivated by the human vision system in a sense that a human describes color by means of hue, saturation and brightness. Hue

and saturation define chrominance, while intensity or value specifies luminance.[8] The HSV color space is defined as follows:[22]

$$\text{Let} \begin{cases} \text{MAX} = \max(R, G, B) \\ \text{MIN} = \min(R, G, B) \\ \delta = \text{MAX} - \text{MIN} \end{cases}$$

$$\begin{aligned} V &= \text{MAX} \\ S &= \begin{cases} \frac{\delta}{\text{MAX}} & \text{if MAX} \neq 0 \\ 0 & \text{if MAX} = 0 \end{cases} \\ H &= \begin{cases} 60\left(\frac{G-B}{\delta}\right) & \text{if MAX} = R \\ 60\left(\frac{B-R}{\delta} + 2\right) & \text{if MAX} = G \\ 60\left(\frac{R-G}{\delta} + 4\right) & \text{if MAX} = B \\ \text{not defined} & \text{if MAX} = 0. \end{cases} \end{aligned} \tag{3}$$

In order to confine H within the range of $[0, 360]$,

$$H = H + 360 \quad \text{if } H < 0.$$

The HSI color space is specified as follows[8]:

$$\begin{aligned} I &= (R + G + B)/3 \\ S &= 1 - (1/I)[\min(R, G, B)] \\ H &= \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{otherwise} \end{cases} \end{aligned} \tag{4}$$

where

$$\theta = \cos^{-1}\left\{\frac{\frac{1}{2}[(R-G) + (R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{\frac{1}{2}}}\right\}.$$

Note that in both Eqs. (3) and (4), the R, G, B values are scaled to $[0, 1]$.

3.4. *Video transmission efficiency color spaces*

The YUV and the YIQ color spaces are commonly used in video for transmission efficiency. The YIQ color space is adopted by the NTSC (National Television System Committee) video standard in reference to $RGB\ NTSC$, while the YUV color space is used by the PAL (Phase Alternation by Line) and the SECAM (System Electronique Couleur Avec Memoire). The YUV color space is specified as

follows:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1471 & -0.2888 & 0.4359 \\ 0.6148 & -0.5148 & -0.1000 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{5}$$

The I and Q components are derived from their counterparts, U and V, via a clockwise rotation (33°), and the YIQ color space is defined as follows:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ 0.5957 & -0.2745 & -0.3213 \\ 0.2115 & -0.5226 & 0.3111 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{6}$$

The $YCbCr$ color space was developed as part of the ITU-R Recommendation B.T. 601[2] for digital video standards and television transmissions. It is a scaled and offset version of the YUV color space. In $YCbCr$, the RGB components are separated into luminance (Y), chrominance blue (Cb) and chrominance red (Cr). The Y component has 220 levels ranging from 16 to 235, while the Cb, Cr components have 225 levels ranging from 16 to 240:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.4810 & 128.5530 & 24.9660 \\ -37.7745 & -74.1592 & 111.9337 \\ 111.9581 & -93.7509 & -18.2072 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{7}$$

where the R, G, B values are scaled to $[0, 1]$.

3.5. *CIE uniform color spaces*

The CIE (Commission Internationale de l'Éclairage) perceptually uniform color spaces, such as the $U^*V^*W^*$, the $L^*u^*v^*$, and the $L^*a^*b^*$ color spaces, are defined based on the XYZ tristimulus:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.607 & 0.174 & 0.200 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.116 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{8}$$

Note that the Y component defined here is consistent with the luminance defined in Eqs. (5)–(7). In addition, a chromaticity diagram can be derived via the chromaticity coordinates x, y, which are specified by the X, Y, Z tristimulus. This CIE chromaticity diagram, however, is not perceptually uniform,[13] i.e. areas of the least perceptible differences on the diagram are distorted to ellipses (known as MacAdam ellipses) rather than circles. To overcome such a shortcoming, the CIE uv chromaticity diagram was proposed[13]:

$$\begin{aligned} u &= 4x/(-2x + 12y + 3) \quad \text{or} \quad 4X/(X + 15Y + 3Z) \\ v &= 6y/(-2x + 12y + 3) \quad \text{or} \quad 6Y/(X + 15Y + 3Z). \end{aligned} \tag{9}$$

Based on this uniform chromaticity scale (UCS), a CIE uniform color space $U^*V^*W^*$ was proposed. The W^* component corresponds to luminance, while the U^*, V^* components correspond to chrominance[13]:

$$\begin{aligned} W^* &= \begin{cases} 116\left(\frac{Y}{Y_o}\right)^{\frac{1}{3}} - 16 & \text{if } \frac{Y}{Y_o} > 0.008856 \\ 903.3\left(\frac{Y}{Y_o}\right) & \text{otherwise} \end{cases} \\ U^* &= 13W^*(u - u_o) \\ V^* &= 13W^*(v - v_o) \end{aligned} \tag{10}$$

where u_o and v_o are derived from the reference white stimulus.

Although the CIE-uv diagram is perceptually uniform, it has its own deficiency in representing yellow-red colors as the area of yellow-red in the diagram is relatively small.[3] To improve this deficiency, a new uv diagram was proposed:

$$\begin{aligned} u' &= u \\ v' &= \frac{3}{2}v. \end{aligned} \tag{11}$$

Based on the $u'v'$ coordinate system, two CIE uniform color spaces were defined, namely the CIE-$L^*u^*v^*$ color space and the CIE-$L^*a^*b^*$ color space.[13] The CIE-$L^*u^*v^*$ color space was proposed to supersede the $U^*V^*W^*$ color space:

$$\begin{aligned} L^* &= \begin{cases} 116\left(\frac{Y}{Y_o}\right)^{\frac{1}{3}} - 16 & \text{if } \frac{Y}{Y_o} > 0.008856 \\ 903.3\left(\frac{Y}{Y_o}\right) & \text{otherwise} \end{cases} \\ u^* &= 13L^*(u' - u'_o) \\ v^* &= 13L^*(v' - v'_o) \end{aligned} \tag{12}$$

where the u'_o and v'_o components are derived from the reference white stimulus.

The $L^*a^*b^*$ color space is one of the most commonly used color spaces and is modeled based on the human vision system. The L^* component in the $L^*a^*b^*$ color space corresponds to brightness ranging from 0 (black) to 100 (white), the a^* component corresponds to the measurement of redness (positive values) or greenness (negative values), and the b^* component corresponds to the measurement of yellowness (positive values) or blueness (negative values). The $L^*a^*b^*$ color space is defined based on the XYZ tristimulus:

$$\begin{aligned} L^* &= 116 f\left(\frac{Y}{Y_o}\right) - 16 \\ a^* &= 500\left[f\left(\frac{X}{X_o}\right) - f\left(\frac{Y}{Y_o}\right)\right] \\ b^* &= 200\left[f\left(\frac{Y}{Y_o}\right) - f\left(\frac{Z}{Z_o}\right)\right] \end{aligned} \tag{13}$$

$$\text{where} \quad f(x) = \begin{cases} x^{\frac{1}{3}} & \text{if } x > 0.008856 \\ 7.787x + \frac{16}{116} & \text{otherwise.} \end{cases}$$

4. Principal Component Analysis

PCA is a standard decorrelation technique and following its application one derives an orthogonal projection basis that directly leads to dimensionality reduction and feature extraction. Let $X \in \mathbb{R}^N$ be a random vector representing an image, where N is the dimensionality of the corresponding image space. The vector is formed by concatenating the rows or the columns of the image which has been normalized to have zero mean and unit variance. The covariance matrix of X is defined as follows:

$$\Sigma_X = E\{[X - E(X)][X - E(X)]^t\} \tag{14}$$

where $E(\cdot)$ is the expectation operator, t denotes the transpose operation, and $\Sigma_X \in \mathbb{R}^{N \times N}$. The PCA of a random vector X factorizes the covariance matrix Σ_X into the following form:

$$\Sigma_X = \Phi \Lambda \Phi^t \tag{15}$$

where $\Phi = [\phi_1 \phi_2 \dots \phi_N] \in \mathbb{R}^{N \times N}$ is an orthogonal eigenvector matrix and $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\} \in \mathbb{R}^{N \times N}$ a diagonal eigenvalue matrix with diagonal elements in decreasing order ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$). $\phi_1, \phi_2, \dots, \phi_N$ and $\lambda_1, \lambda_2, \dots, \lambda_N$ are the eigenvectors and the eigenvalues of Σ_X, respectively.

An important property of PCA is decorrelation, i.e. the components of the transformed data, $X' = \Phi^t X$, are decorrelated since the covariance matrix of X' is diagonal, $\Sigma_{X'} = \Lambda$, and the diagonal elements are the variances of the corresponding components. Another important property of PCA is its optimal signal reconstruction in the sense of minimum Mean Square Error (MSE) when only a subset of principal components is used to represent the original signal. Following this property, an immediate application of PCA is the dimensionality reduction:

$$Y = P^t X \tag{16}$$

where $P = [\phi_1 \phi_2 \dots \phi_m]$, $m < N$, and $P \in \mathbb{R}^{N \times m}$. The lower-dimensional vector $Y \in \mathbb{R}^m$ captures the most expressive features of the original data X.

5. Similarity Measures and Classification Rule

After dimensionality reduction by Eq. (16), feature vectors are compared and classified by the nearest neighbor (to the mean) rule using a similarity (distance) measure δ:

$$\delta(\mathcal{Y}, \mathcal{M}_k) = \min_j \delta(\mathcal{Y}, \mathcal{M}_j) \rightarrow \mathcal{Y} \in \omega_k \tag{17}$$

where $\mathcal{Y}$ is a testing feature vector and $\mathcal{M}_k^0, k = 1, 2, \dots, L$ is the mean of the training samples for class ω_k. The testing feature vector, $\mathcal{Y}$, is classified as belonging to the class of the closest mean, $\mathcal{M}_k$, using the similarity measure δ.

The similarity measures used in our experiments include the L_1 similarity measure, δ_{L_1}, the L_2 similarity measure, δ_{L_2}, the Mahalanobis distance measure, δ_{Md},

and the cosine similarity measure, $\delta_{\cos}$, which are defined as follows:

$$\delta_{L_1}(\mathcal{X}, \mathcal{Y}) = \sum_i |\mathcal{X}_i - \mathcal{Y}_i| \tag{18}$$

$$\delta_{L_2}(\mathcal{X}, \mathcal{Y}) = (\mathcal{X} - \mathcal{Y})^t (\mathcal{X} - \mathcal{Y}) \tag{19}$$

$$\delta_{\mathrm{Md}}(\mathcal{X}, \mathcal{Y}) = (\mathcal{X} - \mathcal{Y})^t \Sigma^{-1} (\mathcal{X} - \mathcal{Y}) \tag{20}$$

$$\delta_{\cos}(\mathcal{X}, \mathcal{Y}) = \frac{-\mathcal{X}^t \mathcal{Y}}{\|\mathcal{X}\| \, \|\mathcal{Y}\|} \tag{21}$$

where Σ in Eq. (20) is the covariance matrix, and $\| \cdot \|$ denotes the norm operator. Note that the cosine similarity measure includes a minus sign in Eq. (21), because the nearest neighbor (to the mean) rule of Eq. (17) applies a minimum rather than a maximum similarity measure.[17]

6. Experiments

This section assesses the performance of content-based face image retrieval in the 12 color spaces defined in Sec. 3: RGB, HSV, YUV, $YCbCr$, XYZ, YIQ, $L^*a^*b^*$, $U^*V^*W^*$, $L^*u^*v^*$, $I_1I_2I_3$, HSI, and rgb. Experimental results using 600 FERET color images corresponding to 200 subjects and 456 FRGC (Face Recognition Grand Challenge) color images of 152 subjects show that some color configurations, such as YV in the YUV color space and YI in the YIQ color space, help improve face retrieval performance. The 600 FERET images are from the FERET database,[19] which has become the *de facto* standard for evaluating face retrieval technologies. The images correspond to 200 subjects such that each subject has three images. The 456 FRGC color images are from the FRGC database,[1] and they correspond to 152 subjects such that there are three images for each subject. Among the three images, one is a controlled image with good image quality, and the remaining two are uncontrolled images with challenging image quality in terms of illumination, resolution, blurring, etc.[1]

The face images used in the experiments are normalized to extract facial regions that contain only faces, so that the face retrieval performance is not affected by the factors not related to face, such as hair style. Specifically, the normalization consists of the following procedures: first, manual annotation detects the centers of the eyes; second, rotation and scaling transformations align the centers of the eyes to the predefined locations and fix the interocular distance; finally, a 128×128 region is cropped to extract the face. Figure 1 shows some example images used in our experiments that are already cropped to the size of 128×128. These images are the individual color component images of the 12 color spaces defined in Sec. 3. From left to right, top to bottom, the color spaces are the RGB color space, the HSV color space, the YUV color space, the $YCbCr$ color space, the XYZ color space, the YIQ color space, the $L^*a^*b^*$ color space, the $U^*V^*W^*$ color space, the $L^*u^*v^*$ color space, the $I_1I_2I_3$ color space, the HSI color space, and the rgb color space. Note that both the FERET and the FRGC images were acquired during different

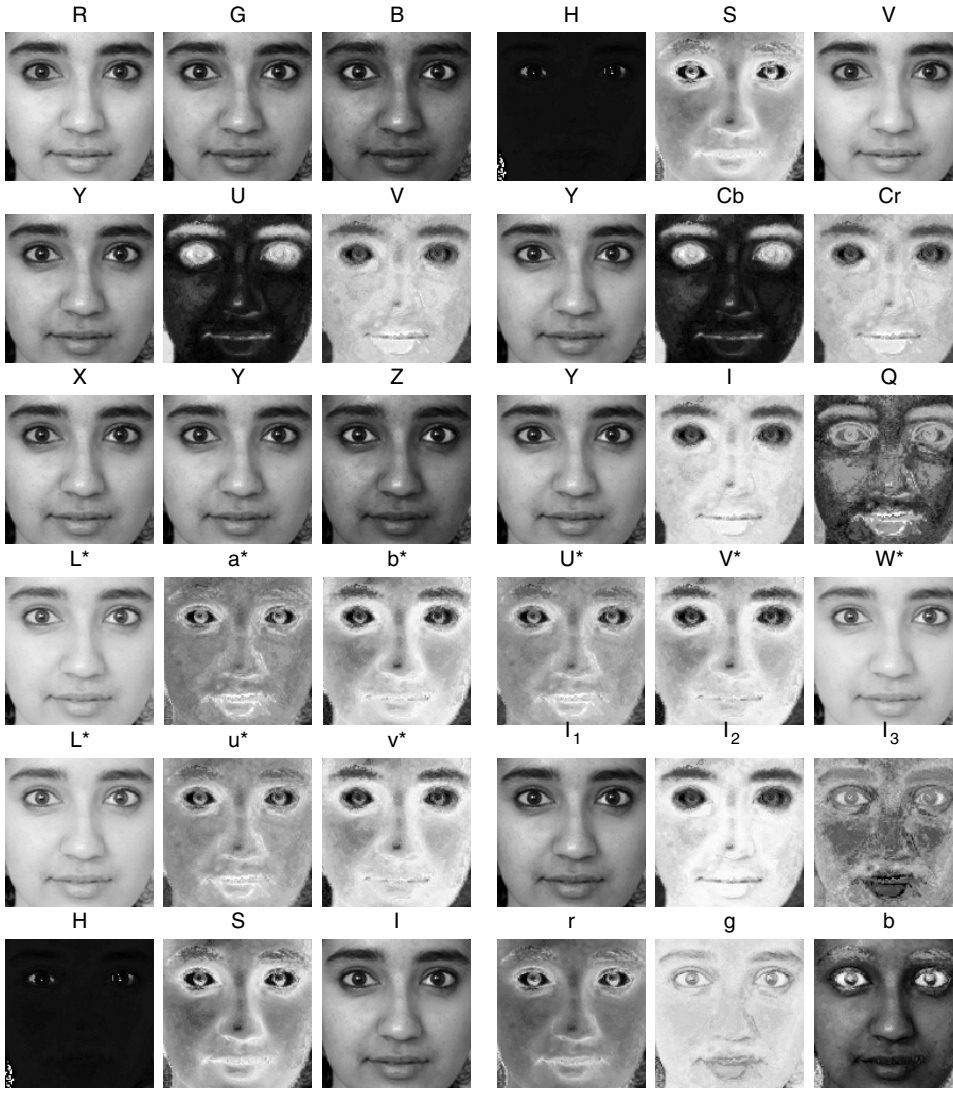

Fig. 1. Color component images of the 12 color spaces as defined in Sec. 3. From left to right, top to bottom, the color spaces are the RGB color space, the HSV color space, the YUV color space, the $YCbCr$ color space, the XYZ color space, the YIQ color space, the $L^*a^*b^*$ color space, and the $U^*V^*W^*$ color space, the $L^*u^*v^*$ color space, the $I_1I_2I_3$ color space, the HSI color space, and the rgb color space.

photo sessions, and thus display different illumination characteristics and facial expressions. For the FERET images, as there are three images of each subject, two images are randomly chosen for training, while the remaining one (unseen during the training) is used for testing; similarly, for the FRGC images, two are randomly chosen for training, while the remaining one is used for testing.

We now comparatively assess the performance of content-based face image retrieval in the 12 color spaces defined in Sec. 3 using the FERET face images. For each color space, we define seven color configurations by means of an individual or a combination of color component images. Take the RGB color space as an example, possible color configurations are R, G, B, RG, RB, GB and RGB. Note that when two or three color component images are used to define a color configuration, each color component image is first normalized to zero mean and unit variance, and then the normalized color component images are concatenated to form an augmented vector representing the color configuration.

To provide a baseline performance for comparison, our first set of experiments applies different similarity measures as defined in Sec. 5 on the intensity images derived by averaging the R, G, B color components. Figure 2 shows the performance of content-based face image retrieval using the PCA method. The horizontal axis indicates the number of features used, and the vertical axis represents the correct retrieval rate, which is the accuracy rate for the top response being correct. Figure 2 shows that the Mahalanobis similarity measure performs the best, followed in order by the L_1 similarity measure, the L_2 distance measure, and the cosine similarity measure. The experimental results provide a baseline face retrieval performance

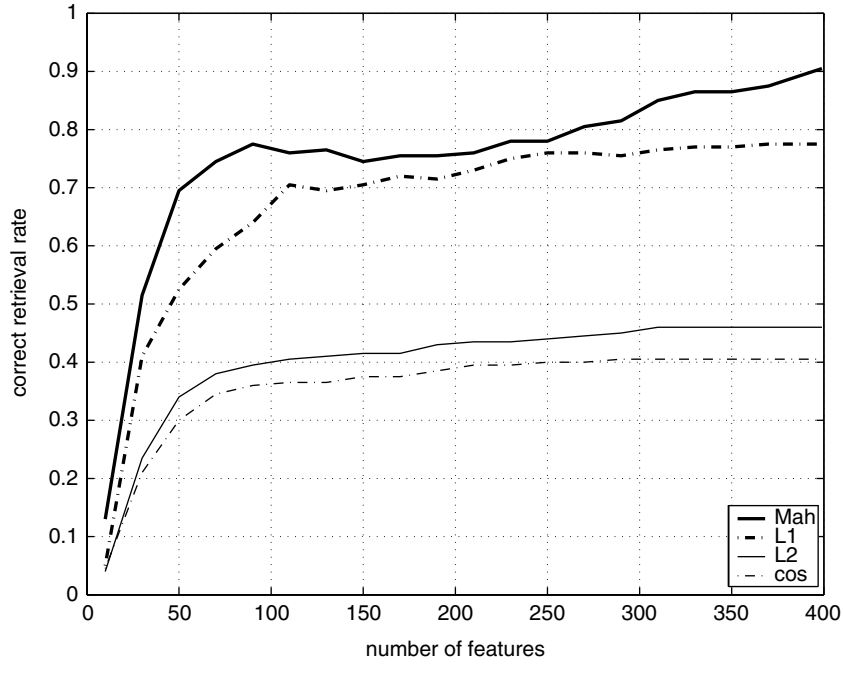

Fig. 2. The performance of content-based face image retrieval using the PCA method on the intensity images derived by averaging the R, G, B color components. The similarity measures applied are the Mahalanobis similarity measure (Mah.), the L_1 similarity measure (L1), the L_2 similarity measure (L2), and the cosine similarity measure (cos).

based on the intensity images, and suggest that one should use the Mahalanobis similarity measure for the comparative assessment in different color spaces. Note that we experimented with all the 12 color spaces as well using these four similarity measures, and experimental results unanimously show that the Mahalanobis similarity measure performs the best.

Our next set of experiments assesses the following color spaces: RGB, HSV, YUV, $YCbCr$, XYZ, YIQ, $L^*a^*b^*$, $U^*V^*W^*$ and $L^*u^*v^*$. The face retrieval performance of the seven color configurations in these color spaces is shown in Figs. 3–9, respectively. Note that the YUV and the $YCbCr$ color spaces (as well as the $U^*V^*W^*$ and the $L^*u^*v^*$ color spaces) have identical face retrieval performance due to their definitions (see Sec. 3). The color configurations, which perform better than the intensity images for face retrieval, are summarized in Table 1. In particular, Fig. 3 shows that the R and the RG color configurations perform better than the intensity images. Figure 4 shows that the V color configuration outperforms the intensity images. Figure 5 shows that the Y and the YV color configurations in the YUV color space or the Y and the YCr color configurations in the $YCbCr$ color space have better face retrieval performance than the intensity images. Figure 6 shows the X, the Y, and the XY color configurations perform better than the intensity images. Figure 7 shows that the Y and the YI color configurations

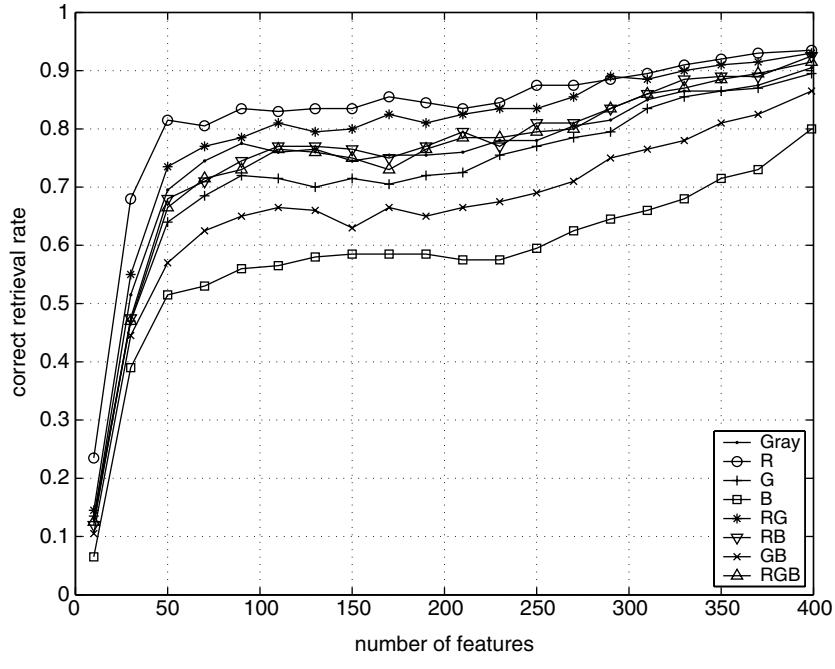

Fig. 3. Content-based face image retrieval performance of the seven color configurations in the *RGB* color space. Note that the performance curve of the intensity images (Gray) is also included for comparison (same in the following figures).

Table 1. The color configurations, which perform better than the intensity images for face retrieval, in the color spaces: RGB, HSV, YUV, $YCbCr$, XYZ, YIQ, $L^*a^*b^*$, $U^*V^*W^*$ and $L^*u^*v^*$.

Color space	Color configurations with better face retrieval performance
RGB	R, RG
HSV	V
YUV	Y, YV
$YCbCr$	Y, YCr
XYZ	X, Y, XY
YIQ	Y, YI
$L^*a^*b^*$	L^*, L^*a^*, L^*b^*, $L^*a^*b^*$
$U^*V^*W^*$	W^*, U^*W^*
$L^*u^*v^*$	L^*, L^*u^*

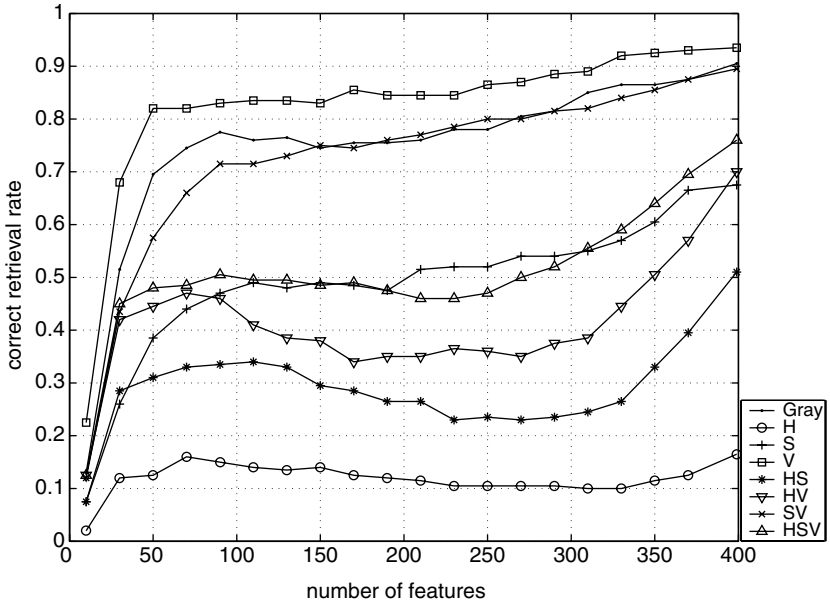

Fig. 4. Content-based face image retrieval performance of the seven color configurations in the HSV color space.

outperform the intensity images. Figure 8 shows that the L^*, the L^*a^*, the L^*b^* and the $L^*a^*b^*$ color configurations are better than the intensity images for face retrieval. Figure 9 shows that the W^* and the U^*W^* color configurations in the $U^*V^*W^*$ color space or the L^* and the L^*u^* color configurations in the $L^*u^*v^*$ color space perform better than the intensity images.

Figure 10 compares the best color configurations in the RGB, HSV, YUV, $YCbCr$, XYZ, YIQ, $L^*a^*b^*$, $U^*V^*W^*$ and $L^*u^*v^*$ color spaces. Note that those

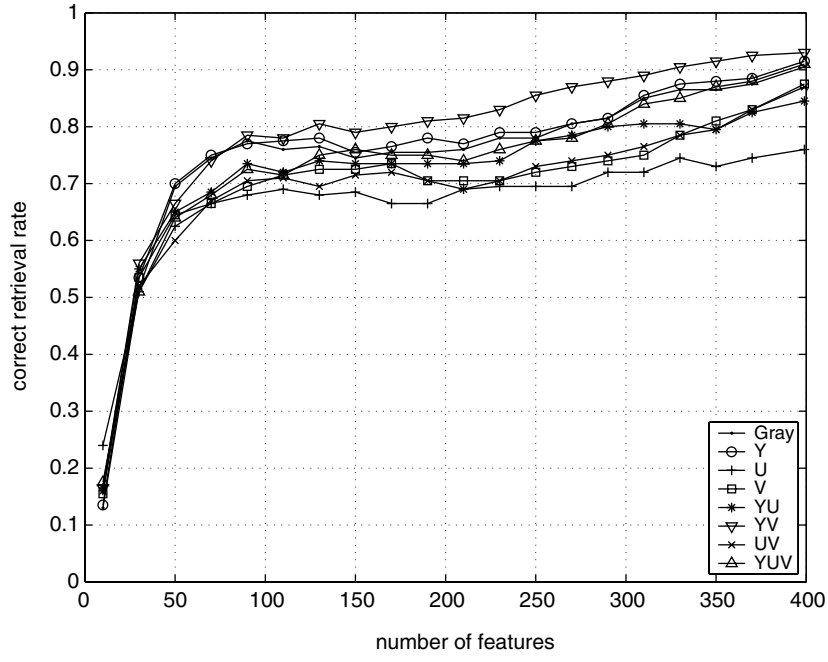

Fig. 5. Content-based face image retrieval performance of the seven color configurations in the YUV color space. Note that the face retrieval results in the $YCbCr$ color space are the same when the Y, U and V color components are replaced by their counterparts Y, Cb and Cr, respectively.

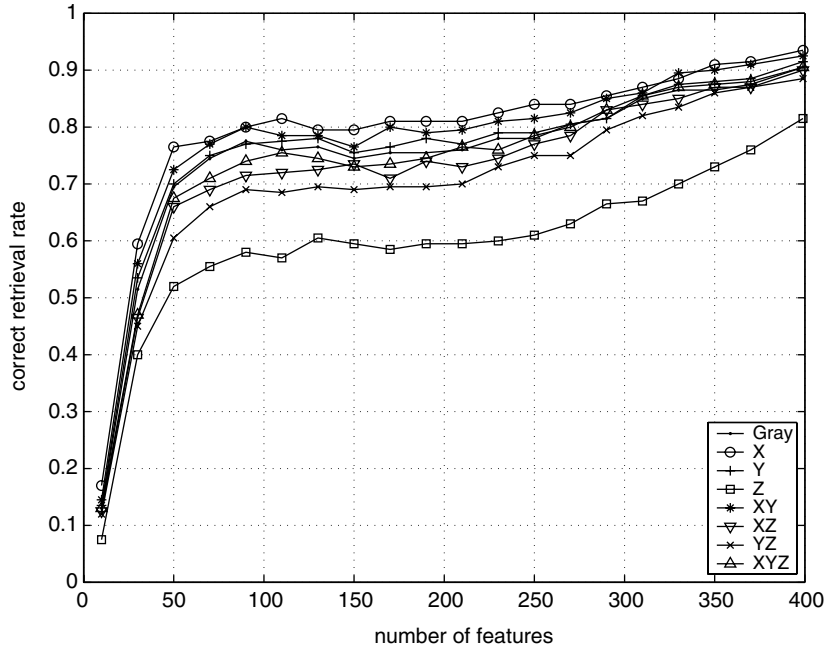

Fig. 6. Content-based face image retrieval performance of the seven color configurations in the XYZ color space.

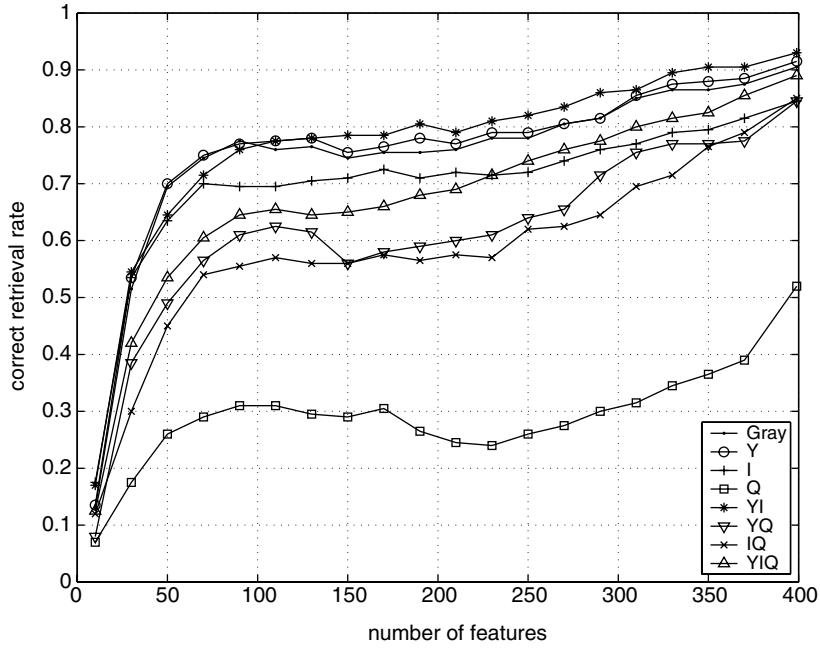

Fig. 7. Content-based face image retrieval performance of the seven color configurations in the YIQ color space.

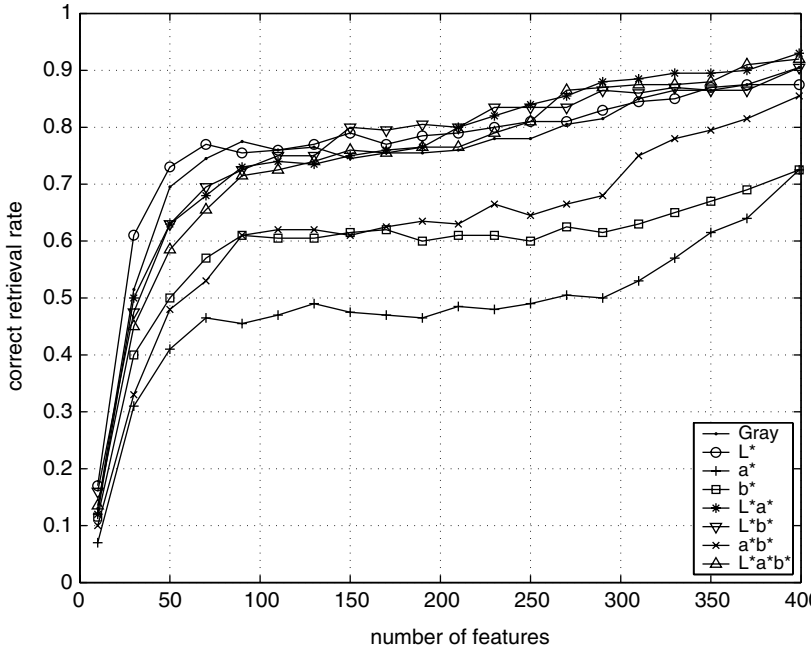

Fig. 8. Content-based face image retrieval performance of the seven color configurations in the $L^*a^*b^*$ color space.

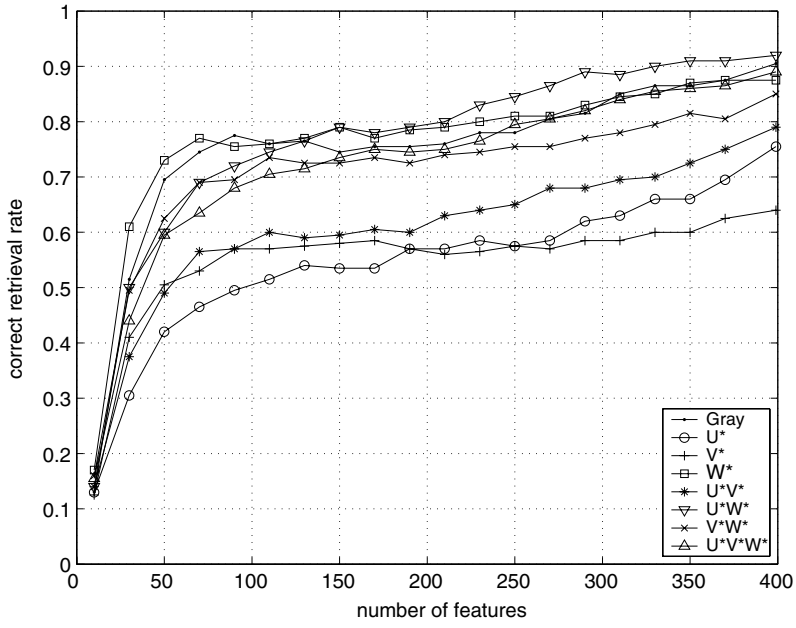

Fig. 9. Content-based face image retrieval performance of the seven color configurations in the $U^*V^*W^*$ color space. Note that the face retrieval results in the $L^*u^*v^*$ color space are the same when the U^*, V^* and W^* color components are replaced by their counterparts u^*, v^* and L^*, respectively.

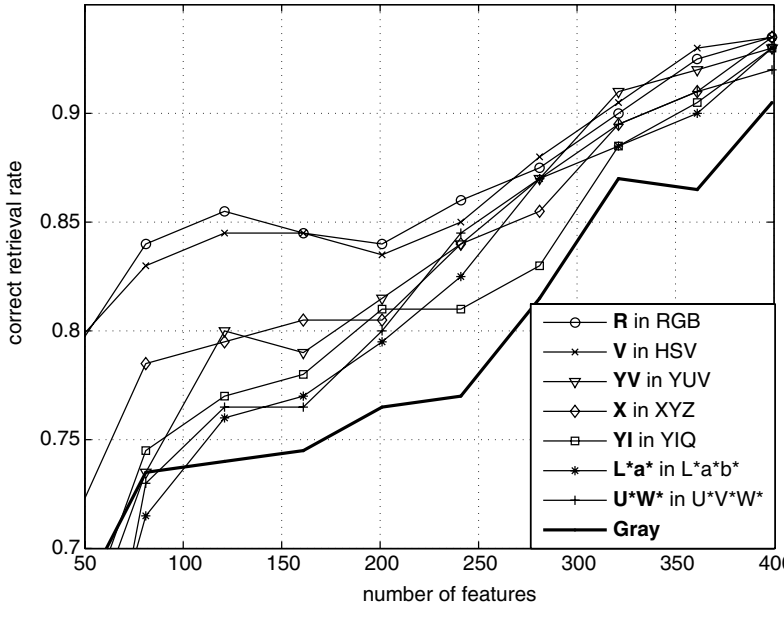

Fig. 10. Comparative content-based face image retrieval performance of the best color configurations in the RGB color space, the HSV color space, the $YUV/YCbCr$ color space, the XYZ color space, the YIQ color space, the $L^*a^*b^*$ color space, and the $U^*V^*W^*/L^*u^*v^*$ color space.

color configurations with better face retrieval performance shown in Fig. 10 all share one common characteristic: they contain both chromatic and achromatic (intensity) components. The pure chromatic color configurations, however, all display worse face retrieval performance than the intensity images. Specifically, these pure chromatic color configurations include the HS in Fig. 4, UV and $CbCr$ in Fig. 5, IQ in Fig. 7, a^*b^* in Fig. 8, and U^*V^* and u^*v^* in Fig. 9. Note also that simply applying all the color components does not necessarily achieve the best face retrieval performance. We have experimented with $I_1I_2I_3$, HSI, and rgb color spaces as well, Figs. 11–13 show that the color configurations in these color spaces do not improve face retrieval performance.

We further assess the robustness of the seven color configurations shown in Fig. 10 using a new dataset from the Face Recognition Grand Challenge (FRGC) database.[1] In particular, the FRGC dataset includes 456 color images of 152 people, and two thirds of the images are taken in uncontrolled environments with challenging image quality in terms of illumination, resolution, blurring, etc.[1] As three images are available for each subject, two of them are randomly chosen for training, while the remaining one is used for testing. Figure 14 shows face retrieval performance of the seven color configurations displayed in Fig. 10. Again, all these seven color configurations have better face retrieval performance than the intensity images, even though the overall face retrieval rates are lower than those shown in

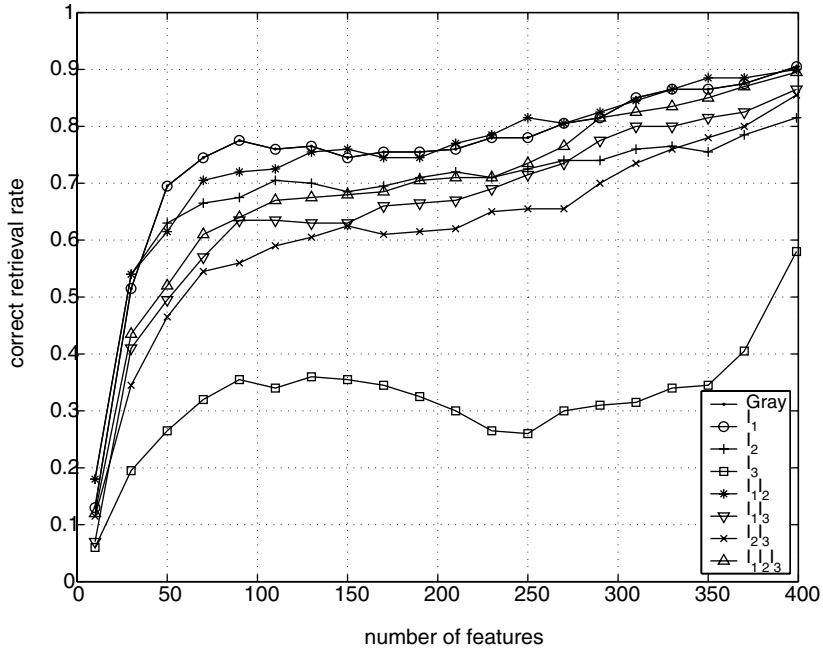

Fig. 11. Content-based face image retrieval performance of the seven color configurations in the $I_1I_2I_3$ color space.

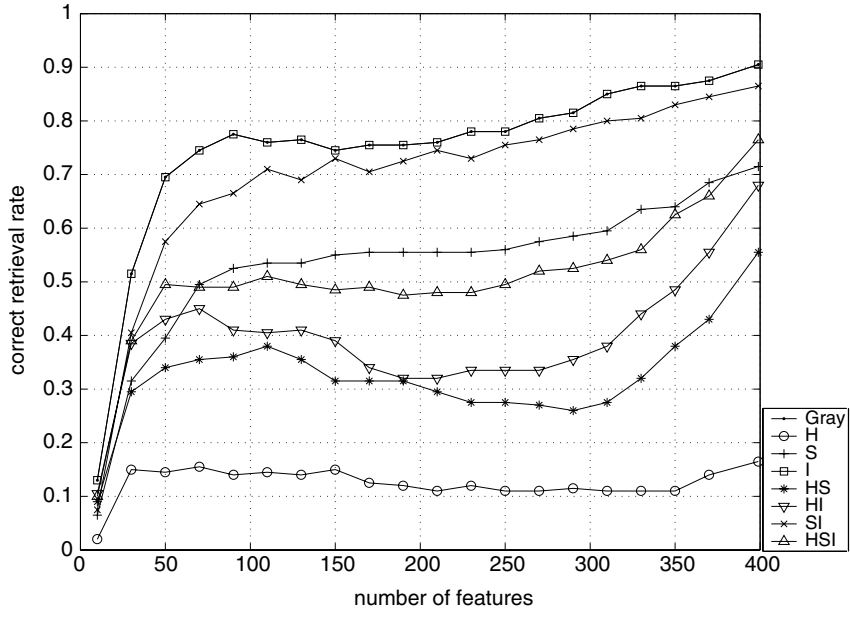

Fig. 12. Content-based face image retrieval performance of the seven color configurations in the HSI color space.

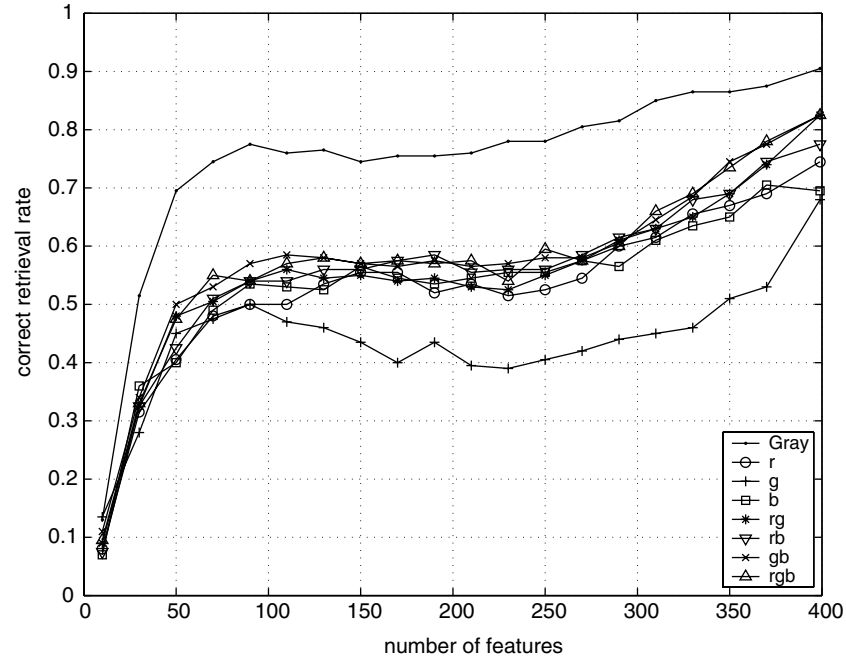

Fig. 13. Content-based face image retrieval performance of the seven color configurations in the rgb color space.

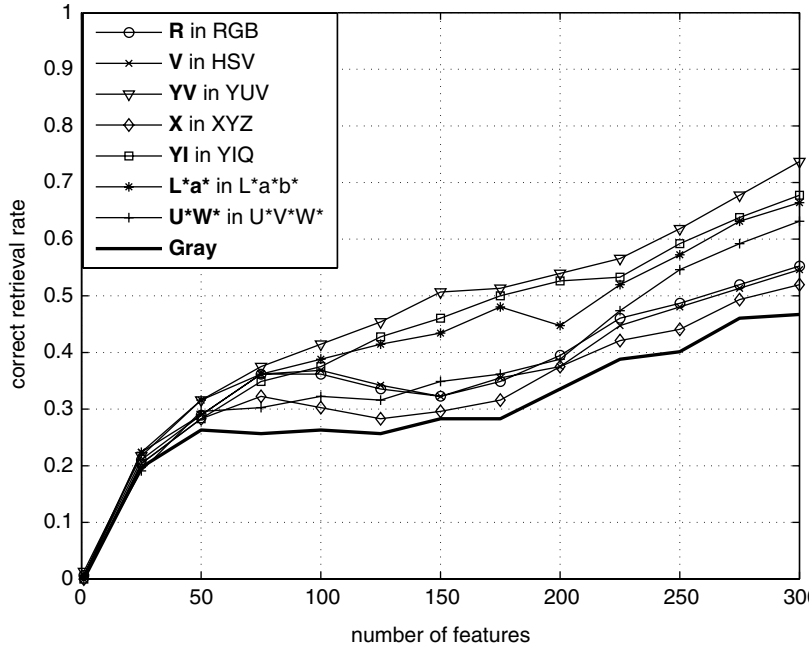

Fig. 14. Comparative content-based face image retrieval performance of seven color configurations using 456 FRGC color images.

Fig. 10. The reason for the lower face retrieval rates is due to the fact that the FRGC dataset contains two thirds of images taken in uncontrolled environments.

We have also experimented with image alignment variations in terms of varying the eye locations to assess the robustness of face retrieval performance using the color configurations shown in Figs. 14 and 10. Our experiments show that (i) small variations in eye locations, such as in 5×5 pixel region, do not change the face retrieval performance much; and (ii) larger variations, such as 11×11 pixel region, can significantly deteriorate the face retrieval performance. Figure 15 shows the deteriorated face retrieval performance when the eye location variations occur in 11×11 pixel region using the FRGC dataset.

7. Conclusion

We have assessed comparatively the performance of content-based face image retrieval in 12 color spaces using a standard algorithm, the PCA, which has become a popular algorithm in the face recognition community. In particular, we have assessed the RGB, HSV, YUV, $YCbCr$, XYZ, YIQ, $L^*a^*b^*$, $U^*V^*W^*$, $L^*u^*v^*$, $I_1I_2I_3$, HSI, and rgb color spaces by evaluating seven color configurations for every single color space. The seven color configurations are defined by an individual or a combination of color component images. Experimental results using 600 FERET

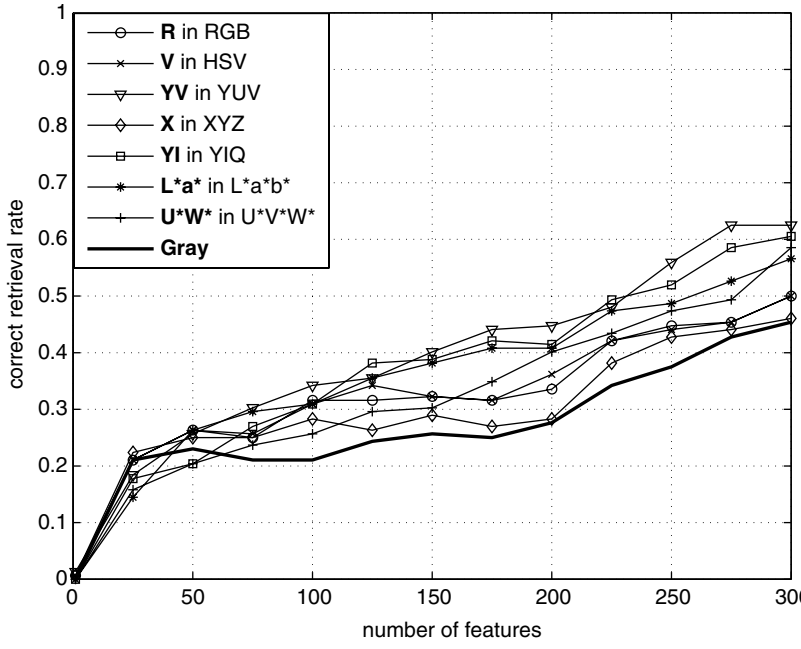

Fig. 15. Comparative content-based face image retrieval performance of seven color configurations using 456 FRGC color images with eye location variations occurred in 11×11 pixel region.

color images corresponding to 200 subjects and 456 FRGC color images of 152 subjects show that some color configurations, such as YV in the YUV color space and YI in the YIQ color space, help improve face retrieval performance.

Acknowledgments

The authors would like to thank the anonymous reviewers for their critical and constructive comments and suggestions. This work was partially supported by the TSWG R&D Contract N41756-03-C-4026.

References

1. K. W. Bowyer, K. Chang and P. Flynn, A survey of approaches to three-dimensional face recognition, in *Proc. 17th Int. Conf. Pattern Recognition* (2004), pp. 358–361.
2. ITU-R Recommendation BT.601-2, Encoding parameters of digital television for studios, in *Int. Telecommunications Union*, Union, Geneva, Switzerland (1982-1986-1990).
3. G. J. Chamberlin and D. G. Chamberlin, *Colour: Its Measurement, Computation and Application* (Heyden & Son, London, 1980).
4. J. Daugman, Face and gesture recognition: overview, *IEEE Trans. Patt. Anal. Mach. Intell.* **19**(7) (1997) 675–676.
5. G. D. Finlayson, S. S. Chatterjee and B. V. Funt, Color angular indexing, in *Proc. European Conf. Computer Vision*, Cambridge, UK (April 14–18, 1996).

6. K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd edn. (Academic Press, 1990).
7. C. Garcia and G. Tziritas, Face detection using quantized skin color regions merging and wavelet packet analysis, *IEEE Trans. Multimed.* **1**(3) (1999) 264–277.
8. R. C. Gonzalez and R. E. Woods, *Digital Image Processing* (Prentice Hall, 2001).
9. N. Habili and C. C. Lim, Hand and face segmentation using motion and color cues in digital image sequences, in *Proc. IEEE Int. Conf. Multimedia and Expo 2001*, Tokyo, Japan (August 22–25, 2001).
10. G. Healey and D. A. Slater, Global color constancy: recognition of objects by use of illumination invariant properties of color distributions, *J. Opt. Soc. Amer.* **A11**(11) (1994) 3003–3010.
11. E. Hjelmas and B. K. Low, Face detection: a survey, *Comput. Vis. Imag. Underst.* **83** (2001) 236–274.
12. R. L. Hsu, M. Abdel-Mottaleb and A. K. Jain, Face detection in color images, *IEEE Trans. Patt. Anal. Mach. Intell.* **24**(5) (2002) 696–706.
13. D. B. Judd and G. Wyszecki, *Color in Business, Science and Industry* (John Wiley & Sons, Inc., 1975).
14. C. Liu, Gabor-based kernel PCA with fractional power polynomial models for face recognition, *IEEE Trans. Patt. Anal. Mach. Intell.* **26**(5) (2004) 572–581.
15. C. Liu and H. Wechsler, Robust coding schemes for indexing and retrieval from large face databases, *IEEE Trans. Imag. Process.* **9**(1) (2000) 132–137.
16. L. Lucchese and S. K. Mitra, Unsupervised color image segmentation, in *Proc. IEEE Workshop on Multimedia Signal Processing*, Los Angeles, CA (December 7–9, 1998).
17. H. Moon and P. J. Phillips, Analysis of pca-based face recognition algorithms, in *Empirical Evaluation Techniques in Computer Vision*, eds. K. W. Bowyer and P. J. Phillips (Wiley-IEEE Computer Society, 1998).
18. Y. Ohta, *Knowledge-Based Interpretation of Outdoor Natural Color Scenes* (Pitman Publishing, London, 1985).
19. P. J. Phillips, H. Wechsler, J. Huang and P. Rauss, The FERET database and evaluation procedure for face-recognition algorithms, *Imag. Vis. Comput.* **16** (1998) 295–306.
20. E. Saber, A. M. Tekalp, R. Eschbach and K. Knox, Automatic image annotation using adaptive color classification, *Graph. Mod. Imag. Process.* **58**(2) (1996) 115–126.
21. A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain, Content-based image retrieval at the end of the early years, *IEEE Trans. Patt. Anal. Mach. Intell.* **22**(12) (2000) 1349–1380.
22. A. R. Smith, Color gamut transform pairs, *Comput. Graph.* **12**(3) (1978) 12–19.
23. K. Sobottka and I. Pitas, Segmentation and tracking of faces in color images, in *Proc. Second Int. Conf. Automatic Face and Gesture Recognition*, Killington, Vermont (October 13–16, 1996).
24. R. Stiefelhagen, J. Yang and A. Waibel, Modeling focus of attention for meeting index based on multiple cues, *IEEE Trans. Neural Networks* **13**(4) (2002) 928–938.
25. M. J. Swain and D. H. Ballard, Color indexing, *Int. J. Comput. Vis.* **7**(1) (1991) 11–32.
26. J. C. Terrillon, M. N. Shirazi, H. Fukamachi and S. Akamatsu, Comparative performance of different skin chrominance models and chrominance space for the automatic detection of human faces in color images, in *Proc. Fourth Int. Conf. Face and Gesture Recognition*, Grenoble, France (March 28–30, 2000).
27. L. Torres, J. Y. Reutter and L. Lorente, The importance of color information in face recognition, in *Proc. IEEE Int. Conf. Image Processing*, Kobe, Japan (October 24–28, 1999).

28. M. Turk and A. Pentland, Eigenfaces for recognition, *J. Cogn. Neurosci.* **13**(1) (1991) 71–86.
29. H. Wu, Q. Chen and M. Yachida, Face detection from color images using a fuzzy pattern matching method, *IEEE Trans. Patt. Anal. Mach. Intell.* **21**(6) (1999) 557–563.
30. M. Yamada, K. Ebihara and J. Ohya, New robust real-time method for extracting human silhouettes from color images, in *Proc. Third Int. Conf. Automatic Face and Gesture Recognition*, Nara, Japan (April 14–16, 1998).
31. J. Yang and A. Waibel, A real-time face tracker, in *Proc. IEEE Workshop on Applications of Computer Vision*, Sarasota, Florida (December 2–4, 1996).
32. H. Yu and J. Yang, A direct lda algorithm for high-dimensional data — with application to face recognition, *Patt. Recogn.* **34**(10) (2001) 2067–2070.
33. W. Zhao and R. Chellappa, Symmetric shape-from-shading using self-ratio image, *Int. J. Comput. Vis.* **45**(1) (2001) 55–75.
34. W. Zhao, R. Chellappa, J. Phillips and A. Rosenfeld, Face recognition: a literature survey, *ACM Comput. Surv.* **35**(4) (2003) 399–458.
35. W. Zhao, A. Krishnaswamy, R. Chellappa, D. L. Swets and J. Weng, Discriminant analysis of principal components for face recognition, in *Face Recognition: From Theory to Applications*, eds. H. Wechsler, P. J. Phillips, V. Bruce, F. F. Soulie and T. S. Huang (Springer-Verlag, 1998), pp. 73–85.

Peichung Shih received his M.S. degree with summa cum laude in information systems in 2002 from New Jersey Institute of Technology, where he is currently a Ph.D. candidate in computer science. His present work includes developing novel face detection systems by applying computer vision concepts and statistical learning theory, and designing robust face recognition models by utilizing color information and genetic computations.

His research interests lie in the field of image and video processing, computer vision, and pattern recognition with applications toward face recognition.

Chengjun Liu received the Ph.D. from George Mason University in 1999, and he is presently an Assistant Professor of Computer Science at New Jersey Institute of Technology.

His recent research has focused on the development of novel and robust methods for image/video retrieval and object detection, tracking and recognition based upon statistical and machine learning concepts.

His research interests are in computer vision, pattern recognition, image processing, evolutionary computation, and neural computation.