Color Face Recognition for Degraded Face Images

Jae Young Choi, Yong Man Ro, *Senior Member, IEEE*, and Konstantinos N. (Kostas) Plataniotis, *Senior Member, IEEE*

***Abstract*—In many current face-recognition (FR) applications, such as video surveillance security and content annotation in a web environment, low-resolution faces are commonly encountered and negatively impact on reliable recognition performance. In particular, the recognition accuracy of current intensity-based FR systems can significantly drop off if the resolution of facial images is smaller than a certain level (e.g., less than 20 × 20 pixels). To cope with low-resolution faces, we demonstrate that facial color cue can significantly improve recognition performance compared with intensity-based features. The contribution of this paper is twofold. First, a new metric called "variation ratio gain" (VRG) is proposed to prove theoretically the significance of color effect on low-resolution faces within well-known subspace FR frameworks; VRG quantitatively characterizes how color features affect the recognition performance with respect to changes in face resolution. Second, we conduct extensive performance evaluation studies to show the effectiveness of color on low-resolution faces. In particular, more than 3000 color facial images of 341 subjects, which are collected from three standard face databases, are used to perform the comparative studies of color effect on face resolutions to be possibly confronted in real-world FR systems. The effectiveness of color on low-resolution faces has successfully been tested on three representative subspace FR methods, including the eigenfaces, the fisherfaces, and the Bayesian. Experimental results show that color features decrease the recognition error rate by at least an order of magnitude over intensity-driven features when low-resolution faces (25 × 25 pixels or less) are applied to three FR methods.**

***Index Terms*—Color face recognition (FR), face resolution, identification, variation ratio gain (VRG), verification (VER), video surveillance, web-based FR.**

I. Introduction

FACE recognition (FR) is becoming popular in research and is being revisited to satisfy increasing demands for video surveillance security [1]–[3], annotation of faces on multimedia contents [4]–[7] (e.g., personal photos and video clips) in web environments, and biometric-based authentication [58]. Despite the recent growth, precise FR is still a challenging task due to ill-conditioned face capturing conditions, such as illumination, pose, aging, and resolution variations between facial images being of the same subject [8]–[10]. In particular, many current FR-based applications (e.g., video-based FR) are commonly confronted with *much-lower-resolution faces* (20 × 20 pixels or less) and suffer largely from them [2], [11], [12]. Fig. 1 shows practical cases in which the faces to be identified or annotated have very small resolutions due to limited acquisition conditions, e.g., faces captured from long distance closed-circuit television (CCTV) cameras or camera phones. As can be seen in Fig. 1(a) and (b), the faces enclosed in red boxes have much lower resolution and additional blurring, which often lead to *unacceptable performance* in the current grayscale (or intensity)-based FR frameworks [13]–[18].

In the practical FR applications, which frequently encounter *low-resolution faces*, it is of utmost importance to select face features that are robust against severe variations in face resolution and to make efficient use of these features. In contrast to the intensity-driven features, color-based features are known to be less susceptible to resolution changes for objection recognition [20]. In particular, the psychophysical results of the FR test in human visual systems showed that the contribution of facial color becomes evident when the shapes of faces are getting degraded [21]. Recently, considerable research effort has been devoted to the efficient utilization of facial color information to improve the recognition performance [22]–[29]. For the color FR reported so far, questions could be categorized as follows: 1) Was color information helpful in improving the recognition accuracy compared with using grayscale only [22]–[29]; 2) how were three different spectral channels of face images incorporated to take advantages of face color characteristics [22], [24], [25], [28], [29]; and 3) which color space was the best for providing discriminate power needed to perform the reliable classification tasks [22], [25], [26]? To our knowledge, however, the color effect on face resolution has not yet been rigorously investigated in the current color-based FR works, and no systematic work suggests the effective color FR framework robust against much-lower-resolution faces in terms of recognition performance.

In this paper, we carry out extensive and systematic studies to explore the facial color effect on the recognition performance as the face resolution is significantly changed. In particular, we demonstrate the significant impact of color on low-resolution faces by comparing the performance between grayscale and color features. The novelty of this paper comes from the following.

1) The derivation of a new metric, which is the so-called variation ratio gain (VRG), for providing the theoretical foundation to prove the significance of color effect on low-resolution faces. Theoretical analysis was made within subspace-based FR methods, which is currently

Manuscript received September 14, 2008; revised December 28, 2008. First published March 24, 2009; current version published September 16, 2009. The work of J. Y. Choi and Y. M. Ro was supported by the Korean Government under Korea Research Foundation Grant KRF-2008-313-D01004. The work of K. N. Plataniotis was supported in part by the Natural Science and Engineering Research Council of Canada under the Strategic Grant BUSNet. This paper was recommended by Associate Editor J. Su.

J. Y. Choi and Y. M. Ro are with the Image and Video System Laboratory, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 305-732, Korea (e-mail: jygchoi@kaist.ac.kr; ymro@ee.kaist.ac.kr).

K. N. Plataniotis is with the Edward S. Rogers, Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada, and also with the School of Computer Science, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: kostas@comm.toronto.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCB.2009.2014245

(a) (b)

Fig. 1. Practical illustrations of extremely small-sized faces in FR-based applications. (a) Surveillance video frame from "Washington Dulles International Airport." The two face regions occupy approximately 18 × 18 pixels of the video frame shown, having an original resolution of 410 × 258 pixels. (b) Personal photo from a "Flickr" [19] web site. The face region occupies about 14 × 14 pixels in the picture shown, having an original resolution of 500 × 333 pixels.

one of the most popular FR techniques [30], [31] due to reliability in performance and simplicity in implementation. VRG quantitatively characterizes how color features affect recognition performance with respect to changes in face resolution.

2) Extensive and comparative recognition performance evaluation experiments to show the effectiveness of color on low-resolution faces. In particular, 3192 frontal facial images corresponding to 341 subjects collected from three public data sets of the Carnegie Mellon University Pose, Illumination, and Expression (CMU PIE) [32], Facial Recognition Technology (FERET) [33], and the Extended Multimodal Verification for Teleservices and Security Applications Database (XM2VTSDB) [34] were used to demonstrate the contribution of color to improved recognition accuracy over various face resolutions commonly encountered from still-image- to video-based real-world FR systems. In addition, the effectiveness of color has successfully been tested on three representative subspace FR methods—principal component analysis [35] (PCA or "eigenfaces"), linear discriminant analysis [8], [36] (LDA or "fisherfaces"), and Bayesian [37] (or "probabilistic eigenspace"). According to experimental results, the effective use of color features drastically reduces the lower bound of face resolution to be reliably recognizable in the computer FR beyond what is possible with intensity-based features.

The rest of this paper is organized as follows. The next section provides background about the low-resolution-face problem in the current FR works. Section III introduces the proposed color FR framework. In Section IV, we first define variation ratio and then make a theoretical analysis to explain the effect of color on variation ratio. In Section V, based on an analysis made in Section IV, VRG is proposed to provide a theoretical insight on the relationship between color effect and face resolutions. Section VI presents the results of extensive experiments performed to demonstrate the effectiveness of color on low-resolution faces. The conclusion is drawn in Section VI.

II. Related Works

In the state-of-the-art FR research, a few works dealt with face-resolution issues. The main concern in these works would be summarized as follows: 1) what is the minimum face resolution to be potentially encountered with the practical applications and to be detectable and recognizable in the computer FR systems [2], [13], [14], [38]–[40] and 2) how do low-resolution faces affect the detection or recognition performances [15]–[17], [40]. In the cutting-edge FR survey literature [2], 15 × 15 pixels is considered to be the minimum face resolution for supporting reliable detection and recognition. The CHIL project [14] reported that normal face resolution in video-based FR is from 10 to 20 pixels in the eye distance. Furthermore, they indicated that the face region is usually 1/16th of commonly used TV recording video frames of resolutions of 320 × 240 pixels. Furthermore, the FR vendor test (FRVT) 2000 [12] studied the effect of face resolution on the recognition performance until the eye distance on the face is as low as 5 × 5 pixels. In the research fields of face detection, 6 × 6 pixels of faces has been reported so far to be the lowest resolution that is feasible for automatic detection [40]. Furthermore, the authors of [39] proposed the face detection algorithm that supports acceptable detection accuracy, even until 11 × 11 pixels.

Several previous works also examined how low-resolution faces impact on recognition performance [15]–[17]. Their works were carried out through intensity-based FR frameworks. They reported that much-lower-resolution faces *significantly degrade* recognition performance in comparison with higher-resolution ones. In [15], face registration and recognition performances were investigated with various face resolutions ranging from 128 × 128 to 8 × 8 pixels. They revealed that face resolutions below 32 × 32 pixels show a considerable decreased recognition performance in PCA and LDA. In [16], face resolutions of 20 × 20 and 10 × 10 pixels dramatically deteriorated recognition performance compared with 40 × 40 pixels in video-based FR systems. Furthermore, the author of [17] reported that the accuracy for face expression recognition is dropped off below 36 × 48 pixels in the neural-network-based recognizer.

Obviously, low-resolution faces impose a significant restriction on current intensity-based FR applications to accomplish *reliability* and *feasibility*. To handle low-resolution-face problems, resolution-enhancement techniques such as "superresolution" [18], [41], [42] are traditional solutions. These techniques usually estimate high-resolution facial images from several low-resolution ones. One critical disadvantage, however, is that the applicability of these techniques is limited to restricted FR domain. This is because they require a sufficient number of

multiple low-resolution images captured from the same identity for the reliable estimation of high-resolution faces. In practice, it is difficult to always support such requirement in practical applications (e.g., the annotation of low-resolution faces on personal photos or snapshot Web images). Another drawback to these approaches is the requirement of a complex framework for the estimation of an image degradation model. It is also computationally demanding for the reconstruction of a high-resolution face image. In this paper, we propose an effective and simple method of using face color features to overcome the low-resolution-face problem. The proposed color FR method improves degraded recognition accuracy, which is caused by low-resolution faces, by a significant margin compared to conventional intensity-based FR frameworks. In addition, contrary to previous resolution-enhancement algorithms, our approach is not only simple in implementation but also guarantees extended applicability to FR applications where only a single color image with a low resolution is available during actual testing operations.

III. Color FR Framework

In this section, we formulate the baseline color FR framework [20] that can make efficient use of facial color features to overcome low-resolution faces. Red–green–blue (RGB) color face images are first converted into another different color space (e.g., YC_bC_r color space). Let I be a color face image generated in the color space conversion process. Then, let s_m be an mth spectral component vector of I (in the form of a column vector by lexicographic ordering of the pixel elements of 2-D spectral images), where $s_m \in \mathbf{R}^{N_m}$ and $\mathbf{R}^{N_m}$ denotes an N_m-dimensional real space. Then, the face vector is defined as the augmentation (or combination) of each spectral component s_m such that $x = [\, s_1^{\mathrm{T}} \quad s_2^{\mathrm{T}} \quad \cdots \quad s_K^{\mathrm{T}} \,]^{\mathrm{T}}$, where $x \in \mathbf{R}^N$, $N = \sum_{m=1}^{K} N_m$, and T represents the transpose operator of the matrix. Note that each s_m should be normalized to zero mean and unit variance prior to their augmentation. Face vector x can be generalized in that, for $K = 1$, the face vector could be defined by grayscale only, while for $K = 3$, it could be defined by a spectral component configuration like YC_bC_r or YQC_r by column order from YC_bC_r and YIQ color spaces.

Most subspace FR methods are separately divided into the training and testing stages. Given a set $\{I_i\}_{i=1}^{M}$ of M color face images, I_i should be first rescaled into the prototype template size to be used for the creation of a corresponding face vector x_i. With a formed training set $\{x_i\}_{i=1}^{M}$ of M face vector samples, the feature subspace is trained and constructed. The rationale behind the feature subspace construction is to find a projection matrix $\Phi = [\, e_1 \quad e_2 \quad \cdots \quad e_F \,]$ by optimizing criteria to get a lower dimensional feature representation $f = \Phi^{\mathrm{T}} x$, where each column vector e_i is a basis vector spanning the feature subspace $\Phi \in \mathbf{R}^{N \times F}$, and $f \in \mathbf{R}^F$. It should be noted that $F \ll N$. For the testing phase, let $\{g_i\}_{i=1}^{G}$ be a gallery (or target) set consisting of G prototype enrolled face vectors of known individuals, where $g_i \in \mathbf{R}^N$. In addition, let p be an unknown face vector to be identified or verified, which is denoted as a probe (or query), where $p \in \mathbf{R}^N$. To perform FR tasks on the probe, g_i $(i = 1, \ldots, G)$ and p are projected onto the feature subspace to get corresponding feature representations such that

$$f_{g_i} = \Phi^{\mathrm{T}} g_i, \qquad f_p = \Phi^{\mathrm{T}} p \tag{1}$$

where $f_{g_i} \in \mathbf{R}^F$ and $f_p \in \mathbf{R}^F$. A nearest-neighbor classifier is then applied to determine the identity of p by finding the smallest distance between f_{g_i} $(i = 1, \ldots, G)$ and f_p in the feature subspace as follows:

$$\ell(p) = \ell(g_{i^*}), \qquad i^* = \arg \min_{i=1}^{G} \|f_{g_i} - f_p\| \tag{2}$$

where $\ell(\cdot)$ returns a class label of face vectors, and $\|\cdot\|$ denotes the distance metric. To exploit why the role of color is getting significant as face resolution is decreased within our baseline color FR framework, a theoretical analysis will be given in the following sections.

IV. Analysis of Color Effect and Face Resolution

Wang and Tang [43] proposed a face difference model that establishes a unified framework of PCA, LDA, and Bayesian FR methods. Based on this model, intra- and extrapersonal variations of feature subspace are critical factors in determining the recognition performance in the three methods. These two parameters are quantitatively well represented by the variation ratio proposed in [44]. Before exploiting the color effect on the recognition performance with respect to changes in face resolution, we begin by introducing the variation ratio and explore how chromaticity components affect the variation ratio within our color FR framework.[1]

A. Variation Ratio

In PCA, covariance matrix C can be computed by using the differences between all possible pairs of two face vectors [43] included in $\{x_i\}_{i=1}^{M}$ such that

$$C = \sum_{i=1}^{M} \sum_{j=1}^{M} (x_i - x_j)(x_i - x_j)^{\mathrm{T}}. \tag{3}$$

Then, C is decomposed into intra- (or within) and extrapersonal (or between class) covariance matrices [43], which are denoted as IC and EC, respectively. IC and EC are defined as

$$\begin{aligned} \mathrm{IC} &= \sum_{l(x_i) = l(x_j)} (x_i - x_j)(x_i - x_j)^{\mathrm{T}} \\ \mathrm{EC} &= \sum_{l(x_i) \neq l(x_j)} (x_i - x_j)(x_i - x_j)^{\mathrm{T}} \end{aligned} \tag{4}$$

where $l(\cdot)$ is a function that returns a class label of x_i as input.

As pointed out in [43], the total variation that resides in the feature subspace is divided into intra- and extrapersonal

[1] In this paper, a theoretical analysis in only the PCA-based color FR framework is given. Our analysis, however, is readily applied to LDA and Bayesian due to the same intrinsic connection of intra- and extrapersonal variations described in [43].

variations related to IC and EC, respectively. From a classification point of view, it is evident that the recognition performance is enhanced as the constructed feature subspace learns and contains a larger variation of EC than that of IC. From this principle, the ratio of extra- to intrapersonal variations can be adopted as an important parameter that reflects the discriminative power of feature space [45]. To define variation ratio, first, let Φ be an eigenvector matrix of C, and then, let Var_Φ (IC) and Var_Φ (EC) be intra- and extrapersonal variations of the feature subspace spanned by Φ, which are computed as [44]

$$\mathrm{Var}_\Phi(\mathrm{IC}) = \mathrm{tr}(\Phi^{\mathrm{T}}\mathrm{IC}\Phi), \qquad \mathrm{Var}_\Phi(\mathrm{EC}) = \mathrm{tr}(\Phi^{\mathrm{T}}\mathrm{EC}\Phi) \quad (5)$$

where $\mathrm{tr}(\cdot)$ is a trace operator of the matrix. Using (5), the variation ratio (J) is defined as

$$J = \frac{\mathrm{Var}_\Phi(\mathrm{EC})}{\mathrm{Var}_\Phi(\mathrm{IC})}. \quad (6)$$

As J increases, a trained feature subspace relatively includes a larger variation of EC in comparison to that of IC. Therefore, J represents a well-discriminative capability of the feature subspace for classification tasks. In (6), the formulation of Var_Φ (IC) and Var_Φ (EC) is similar to that of the J-statistic [55] used in the field of economics. However, it should be pointed out that the metric is used in a novel and quite different way. In particular, the J-statistic has been used as a criterion function to determine the optimal unknown parameter vectors [55], while Var_Φ (IC) and Var_Φ (EC) are used to represent quantitatively the discriminative "effectiveness" of the feature subspace spanned by Φ.

B. Intra- and Extrapersonal Variations in Color FR

In the following section, without loss of generality, we assume that the ith face vector x_i is a configuration of one luminance (s_{i1}) and two different chromaticity components (s_{i2} and s_{i3}) so that $x_i = [\, s_{i1}^{\mathrm{T}} \quad s_{i2}^{\mathrm{T}} \quad s_{i3}^{\mathrm{T}} \,]^{\mathrm{T}}$. By substituting $[\, s_{i1}^{\mathrm{T}} \quad s_{i2}^{\mathrm{T}} \quad s_{i3}^{\mathrm{T}} \,]^{\mathrm{T}}$ into x_i in (3), C is written as

$$C = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \quad (7)$$

where $C_{mn} = \sum_{i=1}^{M}\sum_{j=1}^{M}(s_{im} - s_{jm})(s_{in} - s_{jn})^{\mathrm{T}}$, and $m, n = 1, 2, 3$. As shown in (7), C is a block covariance matrix whose entries are partitioned into covariance or cross-covariance submatrices C_{mn}. For $m = n$, C_{mn} is a covariance submatrix computed from a set $\{s_{im}\}_{i=1}^{M}$; otherwise, for $m \neq n$, C_{mn} is a cross-covariance submatrix computed between $\{s_{im}\}_{i=1}^{M}$ and $\{s_{in}\}_{i=1}^{M}$, where $C_{mn} = C_{nm}^{\mathrm{T}}$. From (4), the IC and EC decompositions of C shown in (7) are represented as

$$\mathrm{IC} = \begin{bmatrix} \mathrm{IC}_{11} & \mathrm{IC}_{12} & \mathrm{IC}_{13} \\ \mathrm{IC}_{21} & \mathrm{IC}_{22} & \mathrm{IC}_{23} \\ \mathrm{IC}_{31} & \mathrm{IC}_{32} & \mathrm{IC}_{33} \end{bmatrix}$$

$$\mathrm{EC} = \begin{bmatrix} \mathrm{EC}_{11} & \mathrm{EC}_{12} & \mathrm{EC}_{13} \\ \mathrm{EC}_{21} & \mathrm{EC}_{22} & \mathrm{EC}_{23} \\ \mathrm{EC}_{31} & \mathrm{EC}_{32} & \mathrm{EC}_{33} \end{bmatrix} \quad (8)$$

where IC_{mn} and EC_{mn} are

$$\mathrm{IC}_{mn} = \sum_{l(x_i)=l(x_j)} (s_{im} - s_{jm})(s_{in} - s_{jn})^{\mathrm{T}}$$
$$\mathrm{EC}_{mn} = \sum_{l(x_i)\neq l(x_j)} (s_{im} - s_{jm})(s_{in} - s_{jn})^{\mathrm{T}}. \quad (9)$$

Like C, IC and EC are also block covariance matrices.

To explore the color effect on variation ratio, we analyze how IC_{mn} and EC_{mn}, which are computed from two different chromaticity components of s_m and s_n $(m, n = 2, 3)$, impact on the construction of variations of IC and EC in (8). By the proof given in the Appendix, trace values of IC and EC can be written as

$$\mathrm{tr}(\mathrm{IC}) = \sum_{m=1}^{3} \mathrm{tr}(I\Lambda_{mm}), \qquad \mathrm{tr}(\mathrm{EC}) = \sum_{m=1}^{3} \mathrm{tr}(E\Lambda_{mm}) \quad (10)$$

where $I\Lambda_{mm}$ and $E\Lambda_{mm}$ are diagonal eigenvalue matrices of IC_{mm} and EC_{mm}, respectively. Using (5) and the cyclic property of the trace operator, the variations of IC and EC are computed as

$$\mathrm{Var}_\Phi(\mathrm{IC}) = \mathrm{tr}(\Phi\Phi^{\mathrm{T}}\mathrm{IC}) = \mathrm{tr}(\mathrm{IC})$$
$$\mathrm{Var}_\Phi(\mathrm{EC}) = \mathrm{tr}(\Phi\Phi^{\mathrm{T}}\mathrm{EC}) = \mathrm{tr}(\mathrm{EC}) \quad (11)$$

where Φ is an eigenvector matrix of C defined in (7). Furthermore, using (5) and the diagonalization of a matrix, the variations of IC_{mm} and EC_{mm} are computed as

$$\mathrm{Var}_{\Phi_{mm}}(\mathrm{IC}_{mm}) = \mathrm{tr}(I\Lambda_{mm})$$
$$\mathrm{Var}_{\Phi_{mm}}(\mathrm{EC}_{mm}) = \mathrm{tr}(E\Lambda_{mm}) \quad (12)$$

where Φ_{mm} is an eigenvector matrix of C_{mm}, and $m = 1, 2, 3$. It should be noted that, in case of $m = 1$, $\mathrm{Var}_{\Phi_{11}}(\mathrm{IC}_{11})$ and $\mathrm{Var}_{\Phi_{11}}(\mathrm{EC}_{11})$ denote intra- and extrapersonal variations calculated from a luminance component of the face vector, while others $(m = 2, 3)$ are corresponding variations computed from two different chromaticity components.

Substituting (11) and (12) into (10), intra- and extrapersonal variations of the feature subspace spanned by Φ can be represented as

$$\mathrm{Var}_\Phi(\mathrm{IC}) = \sum_{m=1}^{3} \mathrm{Var}_{\Phi_{mm}}(\mathrm{IC}_{mm})$$
$$\mathrm{Var}_\Phi(\mathrm{EC}) = \sum_{m=1}^{3} \mathrm{Var}_{\Phi_{mm}}(\mathrm{EC}_{mm}). \quad (13)$$

From (13), we can see that the variation of IC and EC is equal to the summation of the variations of the respective diagonal submatrices of IC_{mm} and EC_{mm}, respectively. This means that Var_Φ (IC) and Var_Φ (EC) are partially decomposed into three independent portions of $\mathrm{Var}_{\Phi_{mm}}(\mathrm{IC}_{mm})$ and $\mathrm{Var}_{\Phi_{mm}}(\mathrm{EC}_{mm})$, where $m = 1, 2, 3$. This confers an important implication about the effect of color on the variation ratio in the color-based FR. Two different chromaticity components can make an independent contribution to construct the intra- and extrapersonal variations in a separate manner with luminance.

Aside from the independent contribution, since each spectral component of skin-tone color has its own inherent characteristics [38], [46], [47], $\mathrm{Var}_{\Phi_{mm}}(\mathrm{IC}_{mm})$ and $\mathrm{Var}_{\Phi_{mm}}(\mathrm{EC}_{mm})$ may differently be changed by practical facial imaging conditions, e.g., illumination and spatial-resolution variations. As a result, intra- and extrapersonal variations in the color-based FR are formed by the composition of variations computed from each spectral component along with different imaging conditions. On the contrary, in the traditional grayscale-based subspace FR, the distribution of the intra- and extrapersonal variations (denoted as $\mathrm{Var}_{\Phi_{11}}(\mathrm{IC}_{11})$ and $\mathrm{Var}_{\Phi_{11}}(\mathrm{EC}_{11})$) in the feature subspace spanned by Φ_{11} is entirely governed by the statistical characteristic of only the luminance component.

C. Color Boosting Effect on Variation Ratio Along With Face Resolution

Now, we make an analysis of the color effect on variation ratio with respect to changes in face resolutions. Our analysis is based on the following two observations: 1) As proven in Section IV-B, each spectral component can contribute in an independent way to construct the intra- and extrapersonal variations of the feature subspace in color-based FR; as described in [54] and [58], such independent impact on evidence fusion usually facilitates a complementary effect between different components for recognition purposes, and 2) the robustness of the color features against variation in terms of face resolution; previous research [20], [48], [49] revealed that chromatic contrast sensitivity is mostly concentrated on low-spatial frequency regions compared to luminance; this means that intrinsic features of face color are even less susceptible to a decrease or variation of the spatial resolution. Considering these two observations, it is reasonable to infer that two chromaticity components can play a supplement role in boosting the decreased variation ratio caused by the loss in the discriminative power of the luminance component arising from low-resolution face images.

To quantize the color boosting effect on variation ratio over changes in the face resolution, we will now derive a simple metric, which is called *variation ratio grain (VRG)*. Using (6) and (12), the variation ratio, which is parameterized by face resolution (γ), for an intensity-based feature subspace is defined as

$$J^{\mathrm{lum}}(\gamma) = \frac{\mathrm{Var}_{\Phi_{11}(\gamma)}\left(\mathrm{EC}_{11}(\gamma)\right)}{\mathrm{Var}_{\Phi_{11}(\gamma)}\left(\mathrm{IC}_{11}(\gamma)\right)}. \tag{14}$$

It should be noted that all terms in (14) are obtained from a training set of intensity facial images having resolution γ. On the other hand, using (13), the variation ratio for a color-augmentation-based feature subspace is defined as

$$\begin{aligned} J^{\mathrm{lum+chrom}}(\gamma) &= \frac{\mathrm{Var}_{\Phi(\gamma)}\left(\mathrm{EC}(\gamma)\right)}{\mathrm{Var}_{\Phi(\gamma)}\left(\mathrm{IC}(\gamma)\right)} \\ &= \frac{\sum_{m=1}^{3} \mathrm{Var}_{\Phi_{mm}(\gamma)}\left(\mathrm{EC}_{mm}(\gamma)\right)}{\sum_{m=1}^{3} \mathrm{Var}_{\Phi_{mm}(\gamma)}\left(\mathrm{IC}_{mm}(\gamma)\right)}. \end{aligned} \tag{15}$$

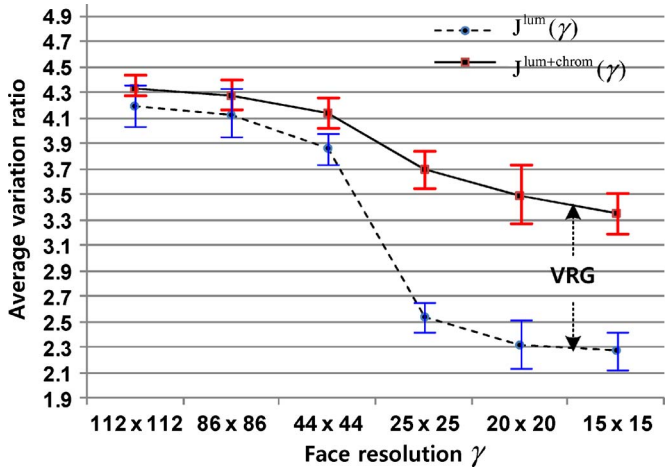

Fig. 2. Average variation ratios and the corresponding standard deviations with respect to six different face-resolution parameters γ. Note that the margin between curves of $J^{\mathrm{lum}}(\gamma)$ and $J^{\mathrm{lum+chrom}}(\gamma)$ represents the numerator of VRG defined as in (16).

Finally, a VRG having input argument γ is defined as

$$VRG(\gamma) = \frac{J^{\mathrm{lum+chrom}}(\gamma) - J^{\mathrm{lum}}(\gamma)}{J^{\mathrm{lum}}(\gamma)} \times 100. \tag{16}$$

$VRG(\gamma)$ measures the relative amount of variation ratio increased by chromaticity components compared to that from only luminance at face resolution γ. Therefore, it reflects well the degree of the effect of color information on the improved recognition performance with respect to changes in γ.

To validate the effectiveness of VRG as a relevant metric for the purpose of quantization of the color effect along with variations in face resolutions, we conducted an experiment using three standard color face DBs of CMU PIE, FERET, and XM2VTSDB. A total of 5000 facial images were collected from three data sets and were manually cropped using the eye position provided by ground truth. Each cropped facial image was first rescaled to a relatively high resolution of 112 $\times$ 112 pixels. To simulate the effect of lowering the face resolution from different distances to the camera, the 5000 facial images with 112 $\times$ 112 pixels were first blurred and then subsequently downsampled by five different factors to produce five different lower-resolution facial images [18]. For blurring, we used a point spread function, which was set to a 5 $\times$ 5 normalized Gaussian kernel with zero mean and a standard deviation of one pixel. After the blurring and downsampling processing, we obtained six sets, each of which consisted of 5000 facial images with six different face resolutions: 112 $\times$ 112, 86 $\times$ 86, 44 $\times$ 44, 25 $\times$ 25, 20 $\times$ 20, and 15 $\times$ 15 pixels (see Fig. 2). We calculated $J^{\mathrm{lum}}(\gamma)$ and $J^{\mathrm{lum+chrom}}(\gamma)$ in (16) over six different face resolution γ parameters. For this, 500 facial images were randomly selected from each set and then used to compute variation ratios by using (14) and (15). The selection process was repeated 20 times so that the variation ratios computed here were the averages of 20 random selections. For luminance and chromaticity components, the YC_bC_r color space was adopted since it has been widely used in image (JPEG) and video (MPEG) compression standards.

Experimental results are shown in Fig. 2. In Fig. 2, $J^{\mathrm{lum}}(\gamma)$ denotes the average variation ratio calculated from luminance face images with resolution γ, i.e., the Y plane from the YC_bC_r color space. Furthermore, $J^{\mathrm{lum+chrom}}(\gamma)$ denotes the average variation ratio computed from YC_bC_r component configuration samples. To guarantee the stability of measured variation ratios, the standard deviations for all cases of $J^{\mathrm{lum}}(\gamma)$ and $J^{\mathrm{lum+chrom}}(\gamma)$ are shown in Fig. 2 as well. As can be seen in Fig. 2, at a high resolution (above 44×44 pixels), the margin between $J^{\mathrm{lum}}(\gamma)$ and $J^{\mathrm{lum+chrom}}(\gamma)$ is relatively small. This is because the luminance component is even more dominant than two chromaticity components in determining $J^{\mathrm{lum+chrom}}(\gamma)$. However, we can observe that $J^{\mathrm{lum}}(\gamma)$ noticeably falls off at low resolution (25×25 pixels or less) compared to those computed from high-resolution faces (above 44×44 pixels). On the other hand, $J^{\mathrm{lum+chrom}}(\gamma)$ has a slower decay compared with $J^{\mathrm{lum}}(\gamma)$ even as the face resolution becomes much lower. In particular, when the face resolution γ is below 25×25 pixels, the difference between $J^{\mathrm{lum}}(\gamma)$ and $J^{\mathrm{lum+chrom}}(\gamma)$ is much larger compared to cases of face resolutions above 44×44 pixels. This result is mostly due to the fact that luminance contrast sensitivity drops off at low spatial frequencies much faster than chromatic contrast sensitivity. Hence, two chromaticity components in (15) can compensate a decreased extrapersonal variation caused by luminance faces with low resolution.

V. Experiments

In the practical FR systems, there are two possible FR approaches to perform FR tasks over lower-resolution probe images [13]. The first method is to prepare multiple training sets of multiresolution facial images and then construct multiple feature subspaces, each of which is charged with a particular face resolution of a probe. An alternative method is that a lower-resolution probe is reconstructed to be matched with the prototype resolution of training and gallery facial images by adopting resolution-enhancement or interpolation techniques. The second method would be appropriate in typical surveillance FR applications in which high-quality training and gallery images are usually employed, but probe images transmitted from surveillance cameras (e.g., CCTV) are often at a low resolution. To demonstrate the effect of color on low-resolution faces in both FR scenarios, two sets of experiments have been carried out in our experimentation. The first experiment is to assess the impact of color on recognition performance with varying face resolutions of probe-given multiresolution trained feature subspaces. On the other hand, the second experiment is to conduct the same assessment when a single-resolution feature subspace trained with high-resolution facial images is only available to the actual testing operation.

A. Face DB for the Experiment and FR Evaluation Protocol

Three de facto standard data sets of CMU PIE, Color FERET, and XM2VTSDB have been used to perform the experiments. The CMU PIE [32] includes 41 368 color images of 68 subjects (21 samples/subject). Among them, 3805 images have the coordinate information of facial feature points. From these

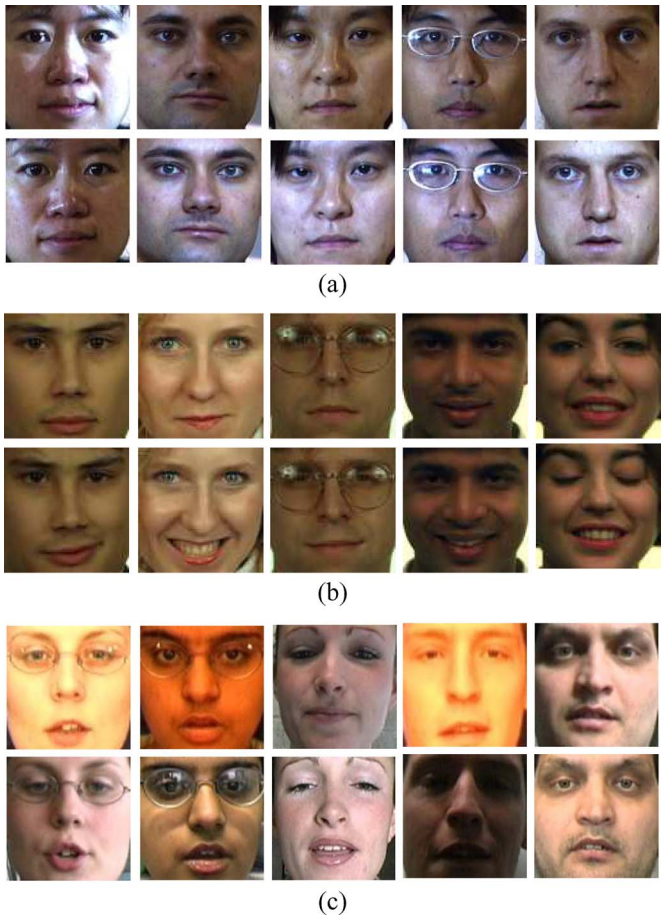

Fig. 3. (a) Examples of facial images from CMU PIE. These images have illumination variations with "room lighting on" conditions. (b) Examples of facial images from FERET. The first and second rows show image examples of **fa** and **fb** sets. (c) Examples of facial images from XM2VTSDB. Note that the facial images in each column belong to the same subject, and all facial images are manually cropped using eye coordinate information. Each cropped facial image is rescaled to the size of 112×112 pixels.

3805 images, 1428 frontal-view facial images with neutral expression and illumination variations were selected in our experimentation. For one subject, 21 facial images have 21 different illumination variations with "room lighting on" conditions. The Color FERET [33] consists of 11 388 facial images corresponding to 994 subjects. Since the facial images are captured over the course of 15 sessions, there are pose, expression, illumination, and resolution variations for one subject. To support the evaluation of recognition performance in various FR scenarios, the Color FERET is to be divided into five sets: "**fa**," "**fb**," "**fc**," "**dup1**," and "**dup2"** partitions [33]. The XM2VTSDB [34] is designed to test realistic and challenging FR with four sessions recorded with no control on severe illumination variations. It is composed of facial images taken on digital video recordings from 295 subjects over a period of one month. Fig. 3 shows examples of facial images selected from three DBs. All facial images shown in Fig. 3 were manually cropped from original images using the eye position provided by a ground truth set.

To construct the training and probe (or test) sets in both sets of experiments, a total of 3192 facial images from 341 subjects were collected from three public data sets. During the collection phase, 1428 frontal-view images of 68 subjects were selected

Fig. 4. Examples of facial images from color FERET according to six different face resolutions. A low-resolution observation below the original 112 × 112 pixels is interpolated using nearest-neighbor interpolation.

from CMU PIE; for one subject, facial images had 21 different lighting variations. From the Color FERET, the 700 frontal-view images of 140 subjects (5 samples/subject) were chosen from the **fa**, **fb**, **fc**, and **dup1** sets. From XM2VTSDB, 1064 frontal-view images of 133 subjects were obtained from two different sessions; each subject included eight facial images that contained illumination and resolution variations. Furthermore, we constructed a gallery set composed of 341 different samples corresponding to 341 different subjects to be identified or verified. Note that, here, gallery images had neutral illumination and expression according to the standard regulation for gallery registration described in [59].

To acquire facial images with varying face resolutions, we carried out resizing over the original collected DB sets. Fig. 4 shows examples of facial images containing face-resolution variations used in our experiments. We took original high-resolution images of faces (shown in the leftmost image of Fig. 4), synthetically blurred them with a Gaussian kernel [41], and then downsampled them so as to simulate a lower resolution effect as closely as possible to practical camera lens. As a result, six different face resolutions of 112 × 112, 86 × 86, 44 × 44, 25 × 25, 20 × 20, and 15 × 15 (pixels) were generated to cover face resolutions that are commonly encountered from practical still-image- to video-based FR applications previously reported in [14]–[16], and [18].

Table I shows the grayscale features, different kinds of color spaces and chromatic features, and spectral component configurations used for our experiments. As shown in Table I, for the grayscale face features, the "R" channel from the RGB color space and the grayscale conversion method proposed in [56] were adopted in our experiments. The R channel of skin-tone color is known to be the best monochrome channel for FR [28], [29]. Moreover, in [56], the $0.85 \cdot R + 0.10 \cdot G + 0.05 \cdot B$ is reported to be an optimal grayscale conversion method for face detection. For the spectral component configuration features, the YC_bC_r, YIQ, and $L^*a^*b^*$ color spaces were used in our experimentation. The YIQ color space defined in the National Television System Committee video standard was adopted. The YC_bC_r color space is scaled and is the offset version of the YUV color space [57]. Moreover, the $L^*a^*b^*$ color space defined in the CIE perceptually uniform color space was used. The detailed description of the used color spaces is given in [57]. As described in [57], the YC_bC_r and YIQ color spaces separate RGB into "luminance" (e.g., Y from the YC_bC_r color space) and "chrominance" (or chromaticity) information (e.g., C_b or C_r from the YC_bC_r color space). In addition, since the $L^*a^*b^*$ color space is based on the CIE XYZ color space [57], it is separated into "luminance" (L^*) and "chromaticity" (a^* and b^*) components. To generate the spectral component configurations depicted in Table I, two different chromaticity components from the used color spaces are combined with a selected grayscale component.

TABLE I
GRAYSCALE FEATURES AND DIFFERENT KINDS OF COLOR SPACES AND SPECTRAL COMPONENT CONFIGURATIONS USED IN OUR EXPERIMENTATION. NOTE THAT THE GRAYSCALE FEATURE IS COMBINED WITH THE CHROMATIC FEATURES TO GENERATE THE SPECTRAL COMPONENT CONFIGURATIONS

Grayscale feature	R [28-29] from RGB color space
	R* = 0.85 · R+0.10 · G+0.05 · B [56] from RGB color space
Chromatic features	C_bC_r [25] fom YC_bC_r(4:4:4) color space
	IQ [25] from YIQ color space
	a*b* [25] from L*a*b* color space
	QC_r [26] from YIQ and YC_bC_r(4:4:4) color spaces
Spectral component configurations	RC_bC_r and R*C_bC_r
	RIQ and R*IQ
	RQC_r and R*QC_r
	Ra*b* and R*a*b*

For FR experiments, all facial images were preprocessed according to the recommendation of the FERET protocol [33] as follows: 1) Color facial images were rotated and scaled so that the centers of eye were placed on the specific pixels; 2) color facial images were rescaled into one of fixed template size among six different spatial resolutions; 3) a standard mask was applied to remove nonface portions; 4) each spectral component of color facial images was separately normalized to have zero mean and unit standard deviations; 5) each spectral image was transformed to a corresponding column vector; and 6) each column vector was used to form a face vector defined in Section III, which covers both grayscale only and spectral component configurations shown in Table I.

To show the *stability* of the significance of color effect on low-resolution faces regardless of FR algorithms, three representative FR methods, which are PCA, Fisher's LDA (FLDA), and Bayesian, were employed. In subspace FR methods, the recognition performance heavily relies on the number of linear subspace dimensions (feature dimension) [50]. Thus, the subspace dimension was carefully chosen and then fixed over six different face resolutions to make a fair comparison of performance. For PCA, the PCA process in FLDA, and Bayesian, a well-known 95% energy capturing rule [50] was adopted to determine subspace dimension. In these experiments, the number of training samples was 1023 facial images so that the subspace dimension was experimentally determined as 200 to satisfy the 95% energy capturing rule. Mahalanobis [51],

Euclidean distances, and "maximum *a posteriori* probability" were used for similarity metrics in PCA, FLDA, and Bayesian, respectively.

In FR tasks, the recognition performance results can be reported for identification and verification (VER). Identification performance is usually plotted on a cumulative match characteristic (CMC) curve [33]. The horizontal axis of the CMC curves is the rank, while the vertical axis is the identification rate. The best found correct recognition rate (BstCRR) [50] was adopted as the identification rate for fair comparison. For the VER performance, the receiver operating characteristic (ROC) [52] curve is popular. The ROC curve plots the face VER rate (FVR) versus the false accept rate (FAR). For an experimental protocol, the collected set of 3192 facial images was randomly partitioned into two sets: training and probe (or test) sets. The training set consisted of (3 samples $\times$ 341 subjects) facial images, with the remaining 2169 facial images for the probe set. There was no overlapping between the two sets for an evaluation of the used FR algorithms' *generalization* performance with regard to the color effect on face resolution. To guarantee the reliability of the evaluation, 20 runs of random partitions were executed, and all of the experimental results reported here were averaged over 20 runs.

B. Experiment 1: To Assess the Impact of Color in Multiresolution Trained Feature Subspace FR Scenario

In experiment 1, it should be noted that the face resolution of each pair of training, gallery, and probe sets were all the same. Since six different face resolutions were used, each feature subspace was trained with a respective set of facial images whose spatial resolution was one of six different kinds. We performed the comparative experiment to compare the recognition performances between the two different grayscale features depicted in Table I. Our experimentation indicates that the R grayscale [28], [29] shows a better performance for most of the face resolutions, as shown in Fig. 4, in the PCA, FLDA, and Bayesian methods. However, the performance difference between the two grayscale configurations is marginal. Thus, R was selected as the grayscale feature of choice for the experiments aiming to the effect of color on low-resolution faces. In addition, "RQC_r" shows the best BstCRR performance of all kinds of spectral component configurations represented in Table I in all face resolutions and the three FR algorithms. This result is consistent with a previous one [26] that reported that "QC_r" is the best chromaticity component in the FR grand challenge DB and evaluation framework [33]. Hence, RQC_r was chosen as a color feature in the following experiments.

Fig. 5 shows the CMC curves for the identification rate (or BstCRR) comparisons between the grayscale and color features with respect to six different face resolutions in the PCA, FLDA, and Bayesian FR methods. As can be seen in CMC curves obtained from the grayscale R feature (in the left side of Fig. 5), the differences in BstCRR between face resolutions of 112 $\times$ 112, 86 $\times$ 86, and 44 $\times$ 44 pixels are relatively marginal in all three FR methods. However, the BstCRRs obtained from a low resolution of 25 $\times$ 25 pixels and below tend to be significantly deteriorated in all three FR methods. For example, for PCA, FLDA, and Bayesian methods, the rank-one BstCRRs (identification rate of top response being correct) decline from 77.20%, 83.69%, and 82.46% to 56.03%, 37.29%, and 62.32%, respectively, as face resolution is reduced from 112 $\times$ 112 to 15 $\times$ 15 pixels.

In case of CMC curves from the RQC_r color feature (on the right side of Fig. 5), we can first observe that color information improves the BstCRR compared with grayscale features over all face resolutions in all three FR algorithms. In particular, it is evident that color features make a substantial enhancement of the identification rate as face resolutions are 25 $\times$ 25 pixels and below. In PCA, 56.03%, 59.81%, and 60.97% of rank-one BstCRRs for 15 $\times$ 15, 20 $\times$ 20, and 25 $\times$ 25 grayscale faces increase to 69.70%, 62.16%, and 75.14%, respectively, by incorporating color feature QC_r. In FLDA, the color feature raises rank-one BstCRRs from 37.29%, 49.72%, and 56.48% to 62.16%, 74.64%, and 77.45% for 15 $\times$ 15, 20 $\times$ 20, and 25 $\times$ 25 face resolutions, respectively. Furthermore, in Bayesian, rank-one BstCRRs increase from 62.23%, 69.17%, and 71.05% to 75.14%, 82.46%, and 84.07% for 15 $\times$ 15, 20 $\times$ 20, and 25 $\times$ 25 face resolutions, respectively.

To demonstrate the color effect on the VER performance according to face-resolution variations, the ROC curves are shown in Fig. 6. We followed the protocol of FRVT [52] to compute the FVR to the corresponding FAR ranging from 0.1% to 100%, and the z-score normalization [54] technique was used. Similar to the identification performance in Fig. 6, face color information significantly improves the VER performance at low-resolution faces (25 $\times$ 25 pixels and below) compared with high-resolution ones. For example, when facial images with a high resolution of 112 $\times$ 112 pixels are applied to PCA, FLDA, and Bayesian, 5.84%, 4.04%, and 2.18% VER enhancements at a FAR of 0.1% are attained from the color feature in PCA, FLDA, and Bayesian methods, respectively. On the other hand, in case of a low resolution of 15 $\times$ 15 pixels, the color feature achieves 19.46%, 38.58%, 15.90% VER improvement at the same FAR for the respective method.

Table II shows the comparison results of VRGs defined in (16) with respect to six different face resolutions in PCA. The $VRG(\gamma)$ for each face resolution γ has been averaged over 20 random selections of 1023 training samples generated from 3192 collected facial images. The corresponding standard deviation for each $VRG(\gamma)$ is also given to guarantee the *stability* of the $VRG(\gamma)$ metric. From Table II, we can see that $VRG(\gamma)$ computed from high-resolution facial images (higher than 44 $\times$ 44 pixels) are relatively small compared with those from low-resolution images (25 $\times$ 25 pixels or lower). This result is largely attributed to the dominance of grayscale information at high-resolution facial images to build intra- and extrapersonal variations in the feature subspace, so that the contribution of color is comparatively small. Meanwhile, in low-resolution color faces, $VRG(\gamma)$ becomes much larger, since color information can *boost* the decreased extrapersonal variation, thanks to its resolution-invariant contrast characteristic and independent impact on constructing variations of feature subspace [20]. The results in Table II verify that face color features play a supplement role in maintaining an extrapersonal variation of feature subspace against face-resolution reduction.

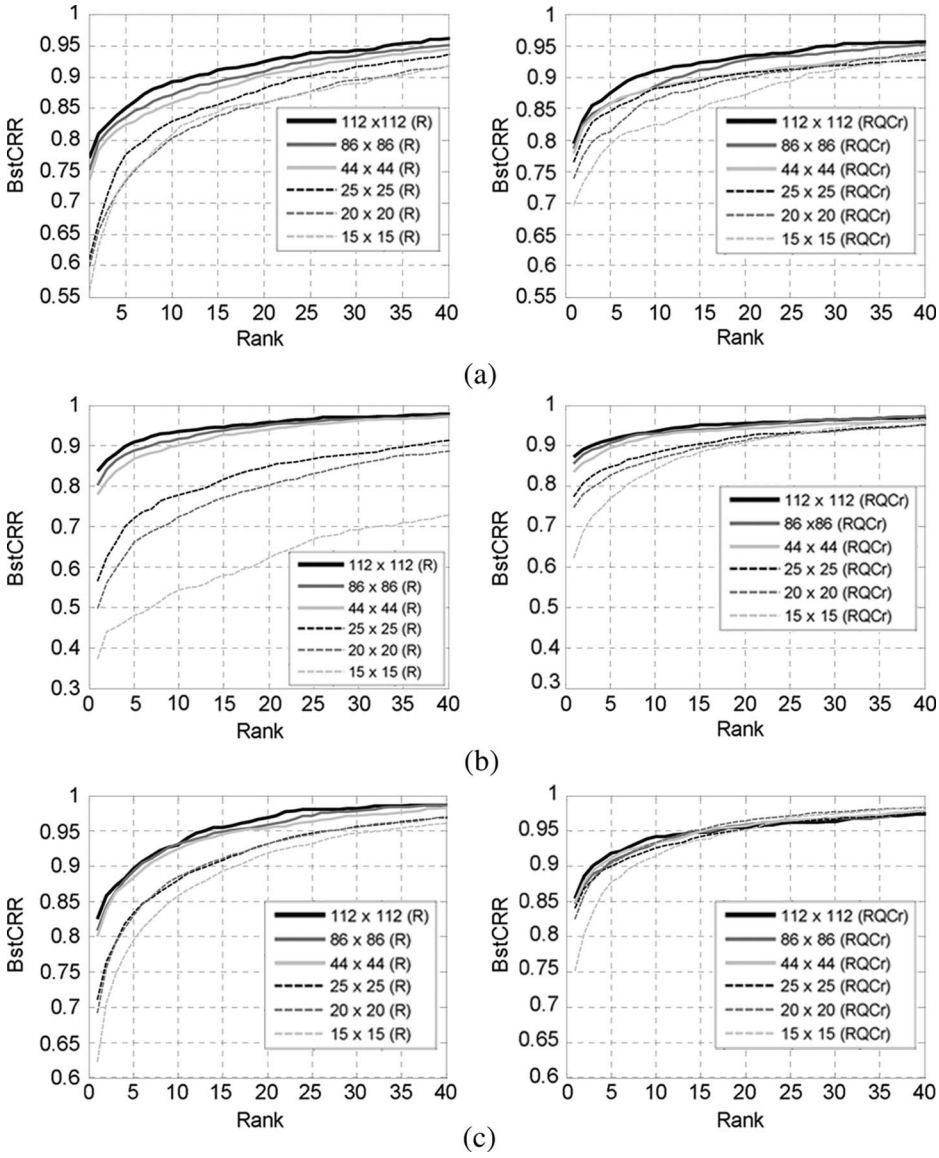

Fig. 5. Identification rate (or BstCRR) comparison between grayscale and color features with respect to six different face resolutions of each pair of training, gallery, and probe facial images in the three FR methods. The graphs on the left side resulted from grayscale feature R, while those on the right side were generated from color feature RQC_r for each face resolution. (a) PCA. (b) FLDA. (c) Bayesian.

C. Experiment 2: To Assess the Impact of Color in a Single-Resolution Trained Feature Subspace FR Scenario

In the practical subspace-based FR applications with face-resolution constraints (e.g., video surveillance), a single feature subspace is usually provided to perform identification or VER tasks on probes. It is reasonable to assume that the feature subspace is pretrained with relatively high-resolution face images [13]. On the other hand, the probes to be tested may have lower and various face resolutions due to heterogeneous acquisition conditions. Therefore, the objective of Experiment 2 is to evaluate the color effect on recognition performance in the FR scenario where high-resolution training images are used to construct a single feature subspace, while probe images have various face resolutions. In Experiment 2, the face resolution of training images was fixed as 112×112 pixels, while the resolution of probe was varied as six different resolutions, as shown in Fig. 4. Since the high-quality gallery images are usually preregistered in FR systems before testing probes [33], we assume that the resolution of gallery is the same as the training facial images, i.e., 112×112 pixels. In Experiment 2, R from the RGB color space was used as a grayscale feature. Due to the best performance from Experiment 1, RQC_r was adopted as a color feature.

Fig. 7 shows the CMC curves with respect to six different probe resolutions in both cases of grayscale (in the left side) and color features (in the right side) in PCA, FLDA, and Bayesian. To obtain a low-dimensional feature representation for a lower face-resolution probe, the probe has been upsampled to have the same resolution of training faces by using a cubic interpolation technique in Fig. 7. From Fig. 7, in case of a grayscale feature, we can see a considerable identification rate degradation in all three FR methods, considering low-resolution (25×25 pixels and below) probes compared with relatively high-resolution counterparts (above 44×44 pixels). In particular, similar to the

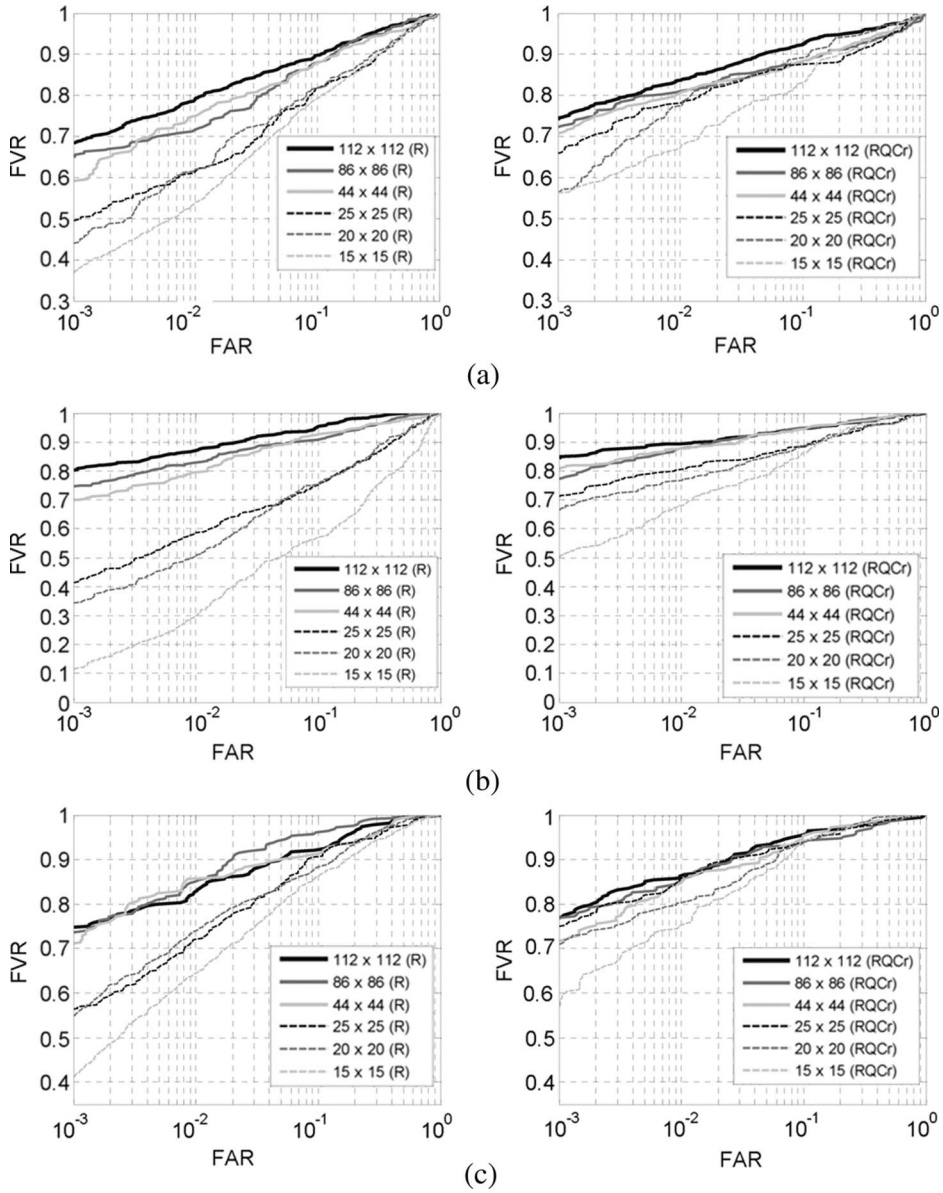

Fig. 6. FVR comparison at FAR ranging from 0.1% to 100% between grayscale and color features with respect to six different face resolutions in the three FR algorithms. The graphs on the left side came from grayscale feature R, while those on the right side were obtained from color feature RQC_r for each face resolution. Note that the z-score normalization technique was used to compute FVR and FAR. (a) PCA. (b) FLDA. (c) Bayesian.

TABLE II
COMPARATIVE EVALUATION OF VRGS DEFINED IN (16) WITH RESPECT TO SIX DIFFERENT FACE RESOLUTIONS OF TRAINING IMAGES IN PCA. GRAYSCALE AND COLOR FEATURES USED FOR COMPUTATION OF VRGS ARE R AND RQC_r (SEE TABLE I), RESPECTIVELY. NOTE THAT THE UNIT OF VRGS IS PERCENT

Face resolution γ	**112 x 112**	**86 x 86**	**44 x 44**	**25 x 25**	**20 x 20**	**15 x 15**
VRG(γ)	4.20%	4.54%	6.80%	51.20%	48.90%	47.20%
Std. Dev.	1.55%	1.78%	2.56%	4.12%	4.20%	3.31%

results from Experiment 1, the identification rate resulting from FLDA is significantly deteriorated at low-resolution probes. The margins of a rank-one identification rate between 112 × 112 and each 25 × 25, 20 × 20, and 15 × 15 pixel grayscale probe in FLDA are 25.66%, 43.77%, and 62.41%, respectively. In case of a color feature, the BstCRR improvement is made at all probe face resolutions in all three FR algorithms. As expected, face color information greatly improves the identification performance obtained from low-resolution probes (25 × 25 pixels and below) compared with grayscale feature. In PCA, by incorporating a color feature, the BstCRR margins between a *grayscale* probe of the 112 × 112 resolution and a *color* probe of the 25 × 25, 20 × 20, and 15 × 15 resolutions are reduced to 3.33%, 4.77%, and 8.02%, respectively. In FLDA, these differences are decreased to 6.65%, 7.28%, and 11.60% at 25 × 25, 20 × 20, and 15 × 15 resolutions, respectively.

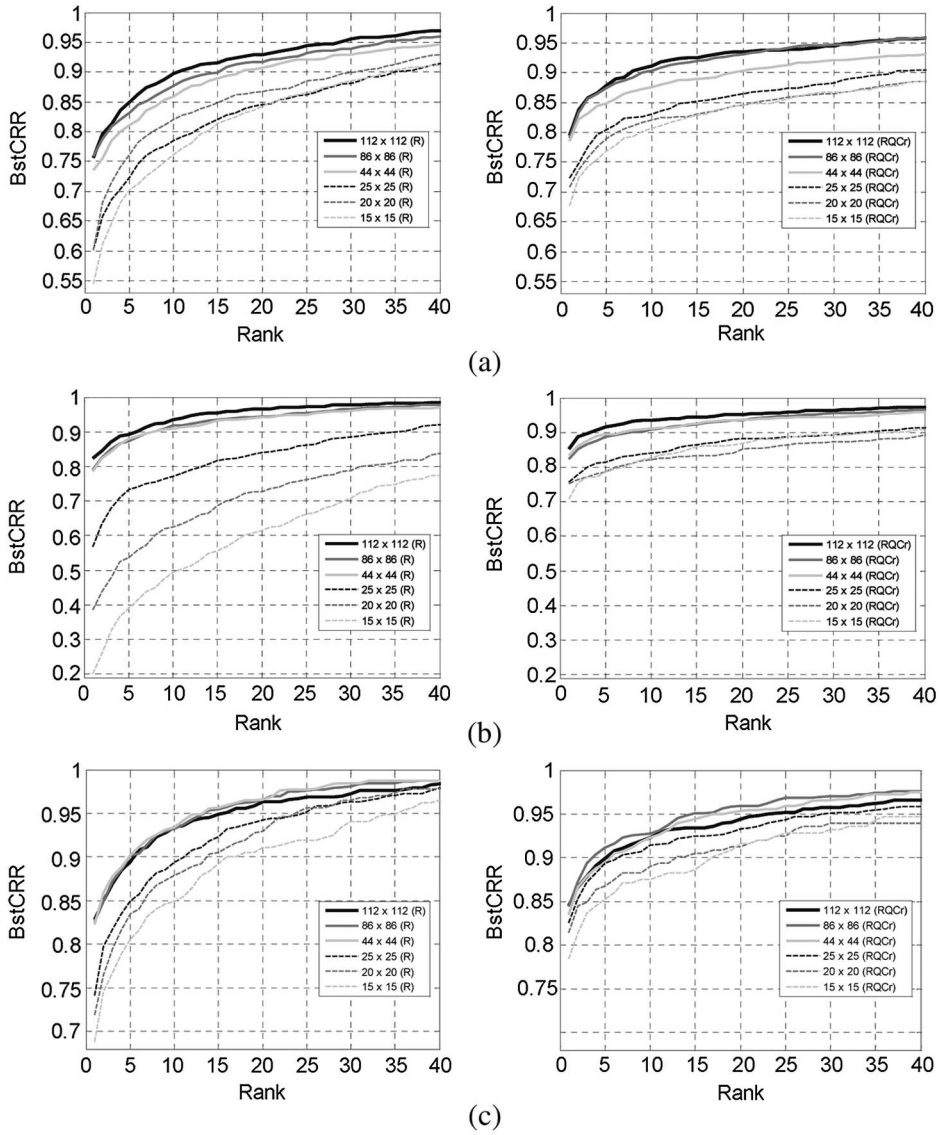

Fig. 7. Identification rate comparison between grayscale and color features with respect to six different face resolutions of probe images. The graphs on the left side resulted from R as a grayscale feature from the RGB color space, while those on the right side were generated from RQC_r as a color feature for each face resolution. Note that a single feature subspace trained with face images having a resolution of 112 × 112 pixels was given to test probe images with varying face resolutions. (a) PCA. (b) FLDA. (c) Bayesian.

In addition, in Bayesian, 1.47%, 2.61%, and 5.64% performance margin decreases are achieved with the aforementioned three different probe resolutions, thanks to the color feature.

Table III presents the FVRs at a FAR of 0.1% obtained from the R grayscale and RQC_r color features with respect to six different face resolutions of probes in three FR methods. Similar to the identification rates shown in Fig. 7, the color feature has a great impact on the FVR improvement at low-resolution faces (25 × 25 pixels and below) in all three FR algorithms. In case of 15 × 15 probe resolutions in PCA, FLDA, and Bayesian, the color feature makes FVR improvements of 15.67%, 54.05%, and 15.62% at a FAR of 0.1%, respectively, in comparison with corresponding FVRs from grayscale probes.

VI. Discussion and Conclusion

According to the results from Experiments 1 and 2, there was a commonly harsh drop-off of identification and VER rates caused by a low-resolution grayscale image (25 × 25 pixels or less) in PCA, FLDA, and Bayesian methods. Considering the performance sensitivity depending on variations in face resolution, FLDA is found to be the weakest to low-resolution grayscale faces (25 × 25 pixels and below) of all three methods. As shown in the CMC curves on the left side of Figs. 5(b) and 7(b), the margins of identification rates between 112 × 112 and 15 × 15 pixels were even 46.40% and 62.41%, respectively. The underlying reason behind such weakness is that optimal criteria used to form the feature subspace in FLDA takes strategy with emphasis on the extrapersonal variation by attempting to maximize it. Therefore, the recognition performance in FLDA is even more sensitive to the portion of extrapersonal variation in the feature subspace compared with the other two methods. Since grayscale features from much-lower-resolution images have a difficulty in providing a sufficient amount of extrapersonal variation to the construction of the feature subspace, the recognition performance could significantly

TABLE III
FVR COMPARISONS AT A FAR OF 0.1% BETWEEN GRAYSCALE AND COLOR FEATURES WITH RESPECT TO SIX DIFFERENT FACE RESOLUTIONS OF PROBE IMAGES IN THE THREE FR ALGORITHMS. R FROM THE RGB COLOR SPACE WAS USED AS A GRAYSCALE FEATURE, WHILE THE RQC_r CONFIGURATION WAS EMPLOYED AS A COLOR FEATURE. NOTE THAT THE z-SCORE NORMALIZATION WAS USED TO COMPUTE FVR VERSUS FAR

FR algorithms (feature)	Face resolution					
	112 x 112	86 x 86	44 x 44	25 x 25	20 x 20	15 x 15
PCA ('R')	68.65%	65.65%	64.01%	41.89%	44.86%	40.01%
PCA ('RQC_r')	72.23%	75.13%	68.37%	62.16%	62.43%	55.67%
FLDA ('R')	78.64%	76.75%	75.94%	32.70%	28.91%	10.81%
FLDA ('RQC_r')	83.51%	80.00%	77.57%	69.45%	69.72%	64.86%
Bayesian ('R')	79.16%	75.19%	76.33%	65.72%	62.72%	52.70%
Bayesian ('RQC_r')	82.06%	78.32%	77.18%	76.56%	74.10%	68.32%

be decreased. On the contrary, thanks to the color's boosting characteristic of the extrapersonal variation, color features in FLDA outperformed by 24.86% and 50.81% margins in case of 15 × 15 pixels, compared with corresponding grayscale images, as shown in Figs. 5(b) and 7(b), respectively. As another interesting finding, Bayesian is more robust to face-resolution variations than PCA and FLDA. For example, from the CMC curves in the left side of Fig. 7, the performance difference between 112 × 112 and 25 × 25 pixels was not so even with 8.54% compared with 15.40% and 25.66% obtained from PCA and FLDA, respectively. A plausible reason under such robustness lies in the fact that Bayesian depends more on the statistical distribution of the intrapersonal variation rather than the extrapersonal variation [30], [37] so that the recognition performance is less likely affected by the reduction of the extrapersonal variation caused by low-resolution images.

Traditionally, low-resolution FR modules have extensively been used in video-surveillance-like applications. Recently, FR applications in the web environment are getting increasing attention due to the popularity of online social networks (e.g., Myspace and Facebook) and their high commercialization potentials [4]–[7]. Under a web-based FR paradigm, many devices such as cellular phone cameras and web cameras often produce low-resolution or low-quality face images which, however, can be used for recognition purposes [4], [5]. As shown in our experimentation, color-based FR outperforms grayscale-based FR over all face resolutions. In particular, thanks to color information, both identification and VER rates obtained by using low-resolution 25 × 25 or 20 × 20 templates are comparable to rates obtained by using much larger grayscale images such as 86 × 86 pixels. Moreover, as shown in Fig. 3, the face DB, which is used in our experimentation, contains images obtained under varying illumination conditions. Hence, the robustness of color in low-resolution FR appears to be stable with respect to the variation in illumination, at least, in our experimentation. These results demonstrate that facial color can reliably and effectively be utilized in real-world FR systems of practical interest, such as video surveillance and promising web applications, which frequently have to deal with low-resolution face images taken under uncontrolled illumination conditions.

APPENDIX

Let $I\Phi_{mm}$ and $I\Lambda_{mm}$ be eigenvector and corresponding diagonal eigenvalue matrices of IC_{mm} in (9), where $m = 1, 2, 3$. That is

$$I\Phi_{mm}^{\mathrm{T}}\mathrm{IC}_{mm}I\Phi_{mm} = I\Lambda_{mm}. \tag{A.1}$$

Using $I\Phi_{mm}(m = 1, 2, 3)$, we define a block diagonal matrix Q given by

$$Q = \mathrm{diag}(I\Phi_{11}, I\Phi_{22}, I\Phi_{33}). \tag{A.2}$$

Note that Q is an orthogonal matrix. Using (8) and (A.2), we now define matrix IS as

$$\begin{aligned}\mathrm{IS} &= Q^{\mathrm{T}}\mathrm{IC\ Q} \\ &= \begin{bmatrix} I\Lambda_{11} & I\Phi_{11}^{\mathrm{T}}\mathrm{IC}_{12}I\Phi_{22} & I\Phi_{11}^{\mathrm{T}}\mathrm{IC}_{13}I\Phi_{33} \\ I\Phi_{22}^{\mathrm{T}}\mathrm{IC}_{21}I\Phi_{11} & I\Lambda_{22} & I\Phi_{22}^{\mathrm{T}}\mathrm{IC}_{23}I\Phi_{33} \\ I\Phi_{33}^{\mathrm{T}}\mathrm{IC}_{31}I\Phi_{11} & I\Phi_{33}^{\mathrm{T}}\mathrm{IC}_{32}I\Phi_{22} & I\Lambda_{33} \end{bmatrix}.\end{aligned} \tag{A.3}$$

IS in (A.3) is *similar* to IC since there exists an invertible matrix Q satisfying $\mathrm{IS} = Q^{-1}\mathrm{ICQ} = Q^{\mathrm{T}}\mathrm{IC}Q$, where $Q^{-1} = Q^{\mathrm{T}}$. Due to their *similarity*, IS and IC have the same eigenvalues and trace value, so that $\mathrm{tr}(\mathrm{IS}) = \mathrm{tr}(\mathrm{IC})$. Note that $\mathrm{tr}(I\Lambda_{mm})$ is the sum of all the eigenvalues of IC_{mm}. Using $\mathrm{tr}(\mathrm{IS}) = \mathrm{tr}(\mathrm{IC})$, $\mathrm{tr}(\mathrm{IC})$ can be expressed as

$$\mathrm{tr}(\mathrm{IC}) = \sum_{m=1}^{3} \mathrm{tr}(I\Lambda_{mm}). \tag{A.4}$$

A similar derivation to (A.1)–(A.3) is also readily applied to EC shown in (8). That is, $\mathrm{tr}(\mathrm{EC})$ can represented as

$$\mathrm{tr}(\mathrm{EC}) = \sum_{m=1}^{3} \mathrm{tr}(E\Lambda_{mm}) \tag{A.5}$$

where $E\Lambda_{mm}(m = 1, 2, 3)$ is a diagonal eigenvalue matrix of EC_{mm}.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their constructive comments and suggestions. The authors would also like to thank the FERET Technical Agent, the U.S. National Institute of Standards and Technology (NIST) for providing the FERET database.

REFERENCES

[1] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," in *Proc. IEEE*, May 1995, vol. 83, pp. 705–740.

[2] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, Dec. 2003.

[3] K. W. Bowyer, "Face recognition technology: Security versus privacy," *IEEE Technol. Soc. Mag.*, vol. 23, no. 1, pp. 9–19, Jun. 2004.

[4] Z. Zhu, S. C. H. Hoi, and M. R. Lyu, "Face annotation using transductive kernel fisher discriminant," *IEEE Trans. Multimedia*, vol. 10, no. 1, pp. 86–96, Jan. 2008.

[5] L. Chen, B. Hu, L. Zhang, M. Li, and H. J. Zhang, "Face annotation for family photo album management," *Int. J. Image Graph.*, vol. 3, no. 1, pp. 1–14, 2003.

[6] S. Satoh, Y. Nakamura, and T. Kanade, "Name-it: Naming and detecting faces in news videos," *IEEE Trans. Multimedia*, vol. 6, no. 1, pp. 22–35, Jan.–Mar. 1999.

[7] J. Y. Choi, S. Yang, Y. M. Ro, and K. N. Plataniotis, "Face annotation for personal photos using context-assisted face recognition," in *Proc. ACM Int. Conf. MIR*, 2008, pp. 44–51.

[8] Z. Wangmeng, D. Zhang, Y. Jian, and W. Kuanquan, "BDPCA plus LDA: A novel fast feature extraction technique for face recognition," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 36, no. 4, pp. 946–953, Aug. 2006.

[9] Q. Li, J. Ye, and C. Kambhamettu, "Linear projection methods in face recognition under unconstrained illumination: A comparative study," in *Proc. IEEE Int. Conf. CVPR*, 2004, pp. II-474–II-481.

[10] R. Singh, M. Vatsa, A. Ross, and A. Noore, "A mosaicing scheme for pose-invariant face recognition," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 37, no. 5, pp. 1212–1225, Oct. 2007.

[11] J. H. Lim and J. S. Jin, "Semantic indexing and retrieval of home photos," in *Proc. IEEE Int. Conf. ICARCV*, 2007, pp. 186–191.

[12] D. M. Blackburn, J. M. Bone, and P. J. Phillips, "Face recognition vendor test 2000: Evaluation report," Defense Adv. Res. Projects Agency, Arlington, VA, 2001.

[13] J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Feature subspace determination in video-based mismatched face recognition," in *Proc. IEEE Int. Conf. AFGR*, 2008, pp. 14–20.

[14] H. K. Ekenel and A. Pnevmatikakis, "Video-based face recognition evaluation in the CHIL project—Run1," in *Proc. IEEE Int. Conf. AFGR*, 2006, pp. 85–90.

[15] B. J. Boom, G. M. Beumer, L. J. Spreeuwers, and R. N. J. Veldhuis, "The effect of image resolution on the performance of a face recognition system," in *Proc. IEEE. Int. Conf. CARV*, 2006, pp. 1–6.

[16] A. Hadid and M. Pietikainen, "From still image to video-based face recognition: An experimental analysis," in *Proc. IEEE Int. Conf. AFGR*, 2004, pp. 813–818.

[17] L. Tian, "Evaluation of face resolution for expression analysis," in *Proc. IEEE Int. Conf. CVPR*, 2004, p. 82.

[18] B. K. Gunturk, A. U. Batur, Y. Altunbasak, M. H. Hayes, III, and R. M. Mersereau, "Eigenface-domain super-resolution for face recognition," *IEEE Trans. Image Process.*, vol. 12, no. 5, pp. 597–606, May 2003.

[19] [Online]. Available: http://www.flickr.com

[20] L. H. Wurm, G. E. Legge, L. M. Isenberg, and A. Lubeker, "Color improves object recognition in normal and low vision," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 19, no. 4, pp. 899–911, Aug. 1993.

[21] A. Yip and P. Sinha, "Role of color in face recognition," *J. Vis.*, vol. 2, no. 7, p. 596, 2002.

[22] L. Torres, J. Y. Reutter, and L. Lorente, "The importance of the color information in face recognition," in *Proc. IEEE Int. Conf. ICIP*, 1999, pp. 627–631.

[23] M. Rajapakse, J. Tan, and J. Rajapakse, "Color channel encoding with NMF for face recognition," in *Proc. IEEE Int. Conf. ICIP*, 2004, vol. 3, pp. 2007–2010.

[24] C. F. Jones, III and A. L. Abbott, "Optimization of color conversion for face recognition," *EURASIP J. Appl. Signal Process.*, vol. 2004, no. 4, pp. 522–529, 2004.

[25] P. Shih and C. Liu, "Comparative assessment of content-based face image retrieval in different color spaces," *Int. J. Pattern Recogn. Artif. Intell.*, vol. 19, no. 7, pp. 873–893, 2005.

[26] P. Shih and C. Liu, "Improving the face recognition grand challenge baseline performance using color configurations across color spaces," in *Proc. IEEE Int. Conf. Image Process.*, 2006, pp. 1001–1004.

[27] B. Karimi, "Comparative analysis of face recognition algorithms and investigation on the significance of color," M.S. thesis, Concordia Univ., Montreal, QC, Canada, 2006.

[28] M. T. Sadeghi, S. Khoushrou, and J. Kittler, "Confidence based gating of colour features for face authentication," in *Proc. Int. Workshop MCS*, 2007, vol. 4472, pp. 121–130.

[29] J. Wang and C. Liu, "A general discriminant model for color face recognition," in *Proc. IEEE Int. Conf. ICCV*, 2007, pp. 1–6.

[30] B. Moghaddam, "Principal manifolds and probabilistic subspaces for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 6, pp. 780–788, Jun. 2002.

[31] J. Lu, K. N. Plataniotis, A. N. Venetsanopoulos, and S. Z. Li, "Ensemble-based discriminant learning with boosting for face recognition," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 166–178, Jan. 2006.

[32] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1615–1618, Dec. 2003.

[33] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, Oct. 2000.

[34] K. Messer, J. Mastas, J. Kittler, J. Luettin, and G. Maitre, "XM2VTSDB: The extended M2VTS database," in *Proc. IEEE Int. Conf. AVBPA*, 1999, pp. 72–77.

[35] M. A. Turk and A. P. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.

[36] P. N. Belhumeur, J. P. Hesphanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern. Anal. Mach. Intell.*, vol. 9, no. 7, pp. 711–720, Jul. 1997.

[37] B. Moghaddam, T. Jebara, and A. Pentland, "Bayesian face recognition," *Pattern Recognit.*, vol. 33, no. 11, pp. 1771–1782, 2000.

[38] R. Hsu, M. A. Monttaleb, and A. Jain, "Face detection in color images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 696–706, May 2002.

[39] A. J. Colmenarez and T. S. Huang, "Face detection and tracking of faces and facial features," in *Proc. IEEE Int. Conf. CVPR*, 1997, pp. 657–661.

[40] S. Hayashi and O. Hasegawa, "A detection technique for degraded face images," in *Proc. IEEE Int. Conf. CVPR*, 2006, pp. 1506–1512.

[41] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 9, pp. 1167–1183, Sep. 2002.

[42] F. W. Wheeler, X. Liu, and P. H. Tu, "Multi-frame super-resolution for face recognition," in *Proc. IEEE Int. Conf. BTAS*, 2007, pp. 1–6.

[43] X. Wang and X. Tang, "A unified framework for subspace face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1222–1228, Sep. 2004.

[44] J. Wang, K. N. Plataniotis, and A. N. Venetasanopoulos, "Selecting discriminant eigenfaces for face recognition," *Pattern Recognit. Lett.*, vol. 26, no. 10, pp. 1470–1482, Jul. 2005.

[45] J. Xiao-Yuan and D. Zhang, "A face and palmprint recognition approach based on discriminant DCT feature extraction," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 6, pp. 2405–2415, Dec. 2004.

[46] H. Stokman and T. Gevers, "Selection and fusion of color models for image feature detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 3, pp. 371–381, Mar. 2007.

[47] Y. Ohta, T. Kanade, and T. Sakai, "Color information for region segmentation," *Comput. Graph. Image Process.*, vol. 13, no. 3, pp. 222–241, Jul. 1980.

[48] D. H. Kelly, "Spatiotemporal variation of chromatic and achromatic contrast thresholds," *J. Opt. Soc. Amer.*, vol. 73, no. 6, pp. 742–749, Jun. 1983.

[49] J. B. Derrico and G. Buchsbaum, "A computational model of spatiochromatic image coding in early vision," *J. Vis. Commun. Image Represent.*, vol. 2, no. 1, pp. 31–38, Mar. 1991.

[50] J. Wang, K. N. Plataniotis, J. Lu, and A. N. Venetsanopoulos, "On solving the face recognition problem with one training sample per subject," *Pattern Recognit.*, vol. 39, no. 6, pp. 1746–1762, Sep. 2006.

[51] V. Perlibakas, "Distance measures for PCA-based face recognition," *Pattern Recognit. Lett.*, vol. 25, no. 12, pp. 1421–1430, Apr. 2004.

[52] P. J. Grother, R. J. Micheals, and P. J. Phillips, "Face recognition vendor test 2002 performance metrics," in *Proc. Int. Conf. Audio- Video-Based Biometric Person Authentication*, 2003, vol. 2688, pp. 937–945.
[53] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, C. Jin, K. Hoffman, J. Marques, M. Jaesik, and W. Worek, "Overview of the face recognition grand challenge," in *Proc. IEEE Int. Conf. CVPR*, 2005, pp. 947–954.
[54] A. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognit.*, vol. 38, no. 12, pp. 2270–2285, Dec. 2005.
[55] L. P. Hansen, "Large sample properties of generalized method of moments estimators," *Econometrica*, vol. 50, no. 4, pp. 1029–1054, 1982.
[56] J. Lu, M. Thiyagarajah, and H. Zhou, "Converting a digital image from color to gray-scale," U.S. Patent 20 080 144 892, Jun. 19, 2008.
[57] R. Lukac and K. N. Plataniotis, *Color Image Processing: Methods and Application*. New York: CRC, 2007.
[58] A. K. Jain, A. Ross, and S. Prabhaker, "An introduction to biometric recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 4–20, Jan. 2004.
[59] *Proposed Draft Amendment to ISO/IEC 19794-5 Face Image Data on Conditions for Taking Pictures*, Mar. 1, 2006.

Jae Young Choi received the B.S. degree from Kwangwoon University, Seoul, Korea, in 2004 and the M.S. degree from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2008, where he is currently working toward the Ph.D. degree with the Image and Video System Laboratory.

He was an Intern Researcher for the Electronic Telecommunications Research Institute, Daejon, in 2007. In 2008, he was a Visiting Student Researcher at the University of Toronto, Toronto, ON, Canada. His research interests include face recognition/detection, image/video indexing, pattern recognition, machine learning, MPEG-7, and personalized broadcasting technologies.

Yong Man Ro (M'92–SM'98) received the B.S. degree from Yonsei University, Seoul, Korea, and the M.S. and Ph.D. degrees from the Korea Advanced Institute in Science and Technology (KAIST), Daejon, Korea.

In 1987, he was a Researcher with Columbia University, New York, NY, and from 1992 to 1995, he was a Visiting Researcher with the University of California, Irvine, and with KAIST. In 1996, he was a Research Fellow with the University of California, Berkeley. He is currently a Professor and the Director of the Image and Video System Laboratory, Korea Advanced Institute of Science and Technology (KAIST), Daejeon. He participated in international standardizations including MPEG-7 and MPEG-21, where he contributed several MPEG-7 and MPEG-21 standardization works, including the MPEG-7 texture descriptor and MPEG-21 DIA visual impairment descriptors and modality conversion. His research interests include image/video processing, multimedia adaptation, visual data mining, image/video indexing, and multimedia security.

Dr. Ro was the recipient of the Young Investigator Finalist Award of the International Society for Magnetic Resonance in Medicine in 1992 and the Scientist Award (Korea), in 2003. He has served as a Technical Program Committee member for many international conferences, including the International Workshop on Digital Watermaking (IWDW), Workshop on Image Analysis for Multimedia Interactive Services (WIAMI), Asia Information Retrieval Symposium (AIRS), Consumer Communications and Networking Conference, etc., and as the Co-Program Chair of the 2004 IWDW.

Konstantinos N. (Kostas) Plataniotis (S'90–M'92–SM'03) received the B.Eng. degree in computer engineering from the University of Patras, Patras, Greece, in 1988 and the M.S. and Ph.D. degrees in electrical engineering from the Florida Institute of Technology, Melbourne, in 1992 and 1994, respectively.

He is currently a Professor with the Edward S. Rogers, Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, where he is a member of the Knowledge Media Design Institute and the Director of Research for the Identity, Privacy, and Security Initiative and is an Adjunct Professor with the School of Computer Science, Ryerson University, Toronto. His research interests include biometrics, communications systems, multimedia systems, and signal and image processing.

Dr. Plataniotis is the Editor-in-Chief for the IEEE SIGNAL PROCESSING LETTERS for 2009–2011. He is a Registered Professional Engineer in the province of Ontario and a member of the Technical Chamber of Greece. He was the 2005 recipient of IEEE Canada's Outstanding Engineering Educator Award "for contributions to engineering education and inspirational guidance of graduate students" and is the corecipient of the 2006 IEEE TRANSACTIONS ON NEURAL NETWORKS Outstanding Paper Award for the paper entitled "Face Recognition Using Kernel Direct Discriminant Analysis Algorithms," which was published in 2003.