

A Color Channel Fusion Approach for Face Recognition

Ze Lu, *Student Member, IEEE*, Xudong Jiang, *Senior Member, IEEE*, and Alex C. Kot, *Fellow, IEEE*

Abstract—Due to high dimensionality of images or generated color features, different color channels are usually processed separately and then concatenated together into a feature vector for classification. This makes channel fusion a crucial step in color face recognition (FR) systems. However, existing methods simply concatenate channel-wise color features without identifying the importance or reliability of features in different color channels. In this paper, we propose a color channel fusion (CCF) approach using jointly dimension reduction algorithms to select more features from reliable and discriminative channels. Experiments using two different dimension reduction approaches, two different types of features on three image datasets show that CCF achieves consistently better performance than color channel concatenation (CCC) method which deals with different color channels equally.

Index Terms—Channel fusion, color face recognition, dimension reduction.

I. INTRODUCTION

COLOR possesses discriminative information for face recognition (FR). Torres *et al.* [1] applied a modified PCA scheme to three color components separately and combined the results. Their results show that the use of color information improves the recognition rate compared to the same scheme using only the luminance information. The improvement can be significant when large facial expression and illumination variations are present or the resolution of face images is low [2], [3]. In [4], a pose invariant face recognition system based on the probability distribution functions of pixels in different color channels was proposed. This method achieved much better performance than that obtained from gray-level face images. Since then, more and more researchers have turned to color information for better FR performance.

Recently, research effort in color FR has been dedicated to applying a multiple-feature encoding scheme to multiple and different color-component images [5]. For example, a discriminative color feature (DCF) method is proposed in [6]. The dimensionality of each color component image is reduced inde-

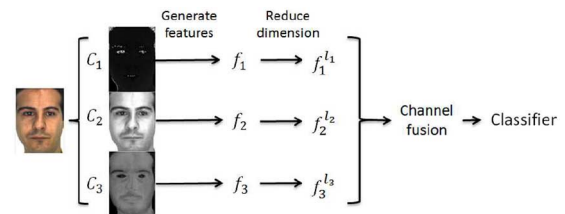


Fig. 1. Color FR framework, C_i , f_i , and $f_i^{l_i}$ indicate color component images, channel-wise features, and low-dimensional features respectively.

pendently and then all low-dimensional color features are concatenated to form an augmented pattern vector. Other than that work, good FR performance is achieved in [5]. The authors propose color local Gabor wavelets (CLGWs) and color LBP (CLBP) features and apply five popular low-dimensional feature extraction techniques including PCA [7] and ERE [8] on each color channel separately. Similar to [9], [10], low-dimensional features are separately extracted from each color component and then combined.

Current color FR works focus on how to extract effective features from multiple color channels following the framework in Fig. 1. Due to the high dimensionality of color component images C_i ($i = 1, 2, 3$) or generated color features f_i , different color channels are usually processed separately first and then concatenated together into a feature vector for classification. According to [11], dimensionality reduction is a critical module of face recognition. However, in current color FR methods, dimensionality reduction is applied on each color channel separately before the channel fusion. Specifically, the dimensionalities of different low-dimensional features ($f_1^{l_1}$, $f_2^{l_2}$, $f_3^{l_3}$) are set to be equal ($l_1 = l_2 = l_3$) in [5], [6], [9], [10]. In fact, the reliability and importance of features in different color channels are not the same, which should be considered in determining l_1, l_2, l_3 . Thus, the rules or ideas of dimension reduction in a single color channel should be integrated across all three color channels to achieve more effective feature extraction and channel fusion. This paper is targeted at filling in this gap by proposing a color channel fusion method where jointly dimension reduction algorithms are used to select more features from reliable and discriminative channels. The main new contribution of this work is selecting more effective features over different color channels for better color channel fusion.

II. COLOR CHANNEL FUSION (CCF) APPROACH

The effectiveness of proposed color channel fusion approach is validated by applying it to two quite different dimensionality reduction methods (PCA and ERE) using two different types of features (image-pixel values and color local Gabor Wavelets

Manuscript received April 15, 2015; revised May 22, 2015; accepted May 22, 2015. Date of publication June 01, 2015; date of current version June 04, 2015. This research was carried out at the Rapid-Rich Object Search (ROSE) Lab at the Nanyang Technological University, Singapore. The ROSE Lab is supported by the National Research Foundation, Prime Ministers Office, Singapore, under its IDM Futures Funding Initiative and administered by the Interactive and Digital Media Programme Office. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Giuseppe Scarpa.

The authors are with the School of Electrical and Electronic Engineering, Nanyang Technological University, 639798 Singapore (e-mail: zlu008@ntu.edu.sg; exdjiang@ntu.edu.sg; eackot@ntu.edu.sg).

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/LSP.2015.2438024

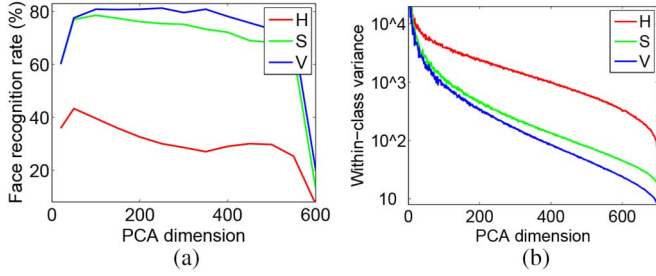


Fig. 2. (a) Channel-wise FR performance against PCA dimension on AR database; (b) within-class variations against PCA dimension on AR database.

(CLGWs) [5]. PCA is commonly used as a benchmark for the evaluation of the performance in FR algorithms [12]. ERE outperforms all other FR methods discussed in [5], [8], [13], [14]. CLGWs is a color local texture feature proposed in [5].

Let I represent the color feature for a face image in one of the color channels H, S, V , then I_{ij} is the feature of j th face image in class i , where $i = 1, 2, \dots, p$, and $j = 1, 2, \dots, q$, so p is the number of subjects, q denotes the number of image samples in class i and there are in total $N = p \times q$ images. Also, \bar{I}_i is the mean feature of samples in class i and \bar{I} is the mean feature across all samples.

A. CCF with PCA

Let's take the dimension reduction process of I_{ij} as an example to explain PCA algorithm, the total scatter matrix of $\{I_{ij}\}$ is defined by

$$\mathbf{S}^t = \frac{1}{N} \sum_{i=1}^p \sum_{j=1}^q (I_{ij} - \bar{I})(I_{ij} - \bar{I})^T. \quad (1)$$

By solving the eigenvalue problem below, eigenvector matrix Φ of \mathbf{S}^t is calculated and Λ is the diagonal matrix of eigenvalues

$$\Lambda = \Phi^T \mathbf{S}^t \Phi. \quad (2)$$

Transformation matrix Φ_l is formed by eigenvectors in Φ corresponding to the l ($l \leq \text{rank of } \mathbf{S}^t$) largest eigenvalues, and I_{ij}^l is the resulting l -dimensional feature vector

$$I_{ij}^l = \Phi_l^T I_{ij}. \quad (3)$$

As the experimental results shown in Fig. 2(a), different color channels possess significantly different classification abilities for face recognition when PCA is used for dimensionality reduction. In other words, the reliability of features in different color channels is not the same.

The question is how to identify the reliability of features in different color channels. As analyzed in [11], [18], [19], PCA improves the generalization capability by removing unreliable dimensions caused by the biased estimates of the within-class variations. The bias is most pronounced when eigenvalues tend toward equality [11], [18], [19]. Fig. 2(b) shows that the channel with larger within-class variations has flatter within-class variation spectrum, so this channel tends to be more unreliable. To apply this principle to the color channel fusion, more dimensions should be selected from the color channel whose within-class variations are smaller.

To derive within-class variation spectrum $V^w \in \mathbb{R}^{1 \times r}$ (r is the rank of \mathbf{S}^t in (1)), within-class scatter matrix is calculated first

$$\mathbf{S}^w = \frac{1}{N} \sum_{i=1}^p \sum_{j=1}^q (I_{ij} - \bar{I}_i)(I_{ij} - \bar{I}_i)^T. \quad (4)$$

Within-class variation along l -th ($l = 1, 2, \dots, r$) eigenvector $\Phi(:, l)$ of \mathbf{S}^t in (2) is

$$V^w(l) = \Phi(:, l)^T \mathbf{S}^w \Phi(:, l). \quad (5)$$

Within-class variation spectrums $V_H^w, V_S^w, V_V^w \in \mathbb{R}^{1 \times r}$ of H, S, V channels can be calculated in the same way. Then the following algorithm is used to determine l_H, l_S, l_V , which indicate the numbers of eigenvectors corresponding to largest eigenvalues used in (3) on H, S, V color channels respectively. Suppose $D = l_H + l_S + l_V$ is the number of dimensions required for the fused feature vector.

$$l_H = 1, l_S = 1, l_V = 1$$

for $n = 4$ to D **do**

$$U_H = \sum_{i=1}^{l_H} V_H^w(i)$$

$$U_S = \sum_{i=1}^{l_S} V_S^w(i)$$

$$U_V = \sum_{i=1}^{l_V} V_V^w(i)$$

$$U_{min} = \min\{U_H, U_S, U_V\}$$

if $U_{min} = U_H$ **then** $l_H = l_H + 1$

else if $U_{min} = U_S$ **then** $l_S = l_S + 1$

else $l_V = l_V + 1$

end if

end for

Initially, l_H, l_S, l_V are set to be 1. In each round of the loop from 4 to D , the channel with smallest sum of within-class variations in selected dimensions is chosen and in this channel, the eigenvector of \mathbf{S}^t corresponding to the largest eigenvalue among the unselected ones is chosen. In this way, more dimensions from more reliable channels across H, S, V are chosen. It is easy to see that this algorithm also tends to make the sum of within-class variations in each color space (U_H, U_S, U_V) equal.

B. CCF with ERE

Besides the dimensionality reduction, regularization is another effective technique to improve the robustness of face recognition [8], [20], [21]. Eigenfeature regularization and extraction (ERE) [8] first regularizes eigenvalues of within-class scatter matrix based on an eigenspectrum model, and then extracts the most discriminant features for fast recognition.

From channel-wise FR performance of different color channels shown in Fig. 3(a), it can be observed that the classification ability of different color channels is not the same when ERE is used for dimensionality reduction. In order to apply ERE's discriminative rule of a single color channel to all three color channels, the criteria based on maximization of discriminant value J

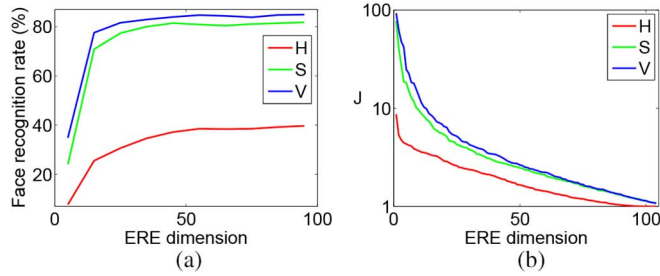


Fig. 3. (a) Channel-wise FR performance against ERE dimension on pose variation subset of MultiPIE; (b) discriminant value J against ERE dimension on pose variation subset of MultiPIE.

used in [8] is adopted. In this way, more features from more discriminant channels across H, S, V are selected.

J spectrums of different color channels are shown in Fig. 3(b). Suppose $\mathbf{A}_H, \mathbf{A}_S, \mathbf{A}_V$ are eigenvector matrices derived by ERE on channel H, S, V respectively. The following algorithm is used to determine l_H, l_S, l_V , which indicate the numbers of eigenvectors in $\mathbf{A}_H, \mathbf{A}_S, \mathbf{A}_V$ corresponding to largest J values used in ERE on H, S, V color channels. Suppose $D = l_H + l_S + l_V$ is the number of dimensions required for the fused feature vector.

$$l_H = 1, l_S = 1, l_V = 1$$

$$J = \{J_H(2 : \text{end}), J_S(2 : \text{end}), J_V(2 : \text{end})\} \text{ arranged in descending order}$$

for $n = 1$ to $D - 3$ **do**

if $J(n) \in J_H$ **then** $l_H \leftarrow l_H + 1$

else if $J(n) \in J_S$ **then** $l_S \leftarrow l_S + 1$

else $l_V \leftarrow l_V + 1$

end if

end for

To begin with, l_H, l_S, l_V are set to be 1. Discriminant value spectrums $J_H(2 : \text{end}), J_S(2 : \text{end}), J_V(2 : \text{end})$ from H, S, V color channels are put together and arranged in descending order. l_H, l_S, l_V are determined by counting how many J values are from corresponding channel among $\{J(n)\}, n = (1, 2, 3, \dots, D - 3)$. Thus, more dimensions from more discriminant channels across H, S, V are selected.

C. Face Recognition

With l_H, l_S, l_V eigenvectors corresponding to largest eigenvalues in PCA or discriminant values in ERE, three transformation matrices are formed to extract low-dimensional features from H, S, V color channels respectively. Then three low-dimensional feature vectors are normalized and concatenated. A minimum-Mahalanobis-distance classifier is subsequently performed between probe and gallery images to determine the identity of probe images.

III. EXPERIMENTS

The proposed color channel fusion (CCF) algorithm is verified by applying it to two quite different dimensionality

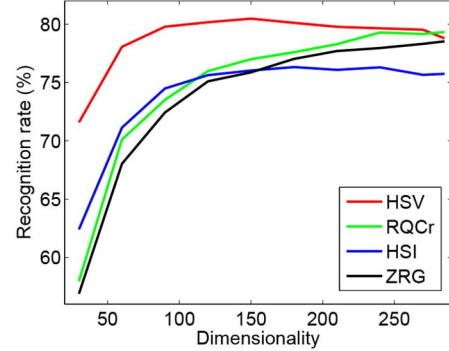


Fig. 4. FR rate of 4 color spaces on AR using ERE-based CCF.

reduction approaches (PCA and ERE) using two different types of features (image-pixel values and CLGWs) on three publicly available datasets: AR [22] and two datasets from CMU Multi-PIE [23]. It is compared with the color channel concatenation (CCC) algorithm used in [4]–[6], [9], [10] and the decision-level fusion method using a weighted sum rule (WDF) in [24].

The face images in the CMU Multi-PIE database are captured under variations of illumination, expression and pose across 4 sessions. The first 105 subjects which appear in all 4 sessions with illumination and pose variations are used in the experiments. Images are cropped based on the eye locations provided in [29]. The AR database contains 2600 frontal-face images captured across 2 sessions from 100 subjects (50 males and 50 females). For the 13 images per subject in each session, 7 undisguised images with mixed variations (expression variation and illumination variation) are used.

A. Color Space Selection

The face recognition performance is not the same in different color spaces. It is well-known that HSV [1], [25], [26], [27], RQCr [3], [5], [10], ZRG [28], [5], [10] and HSI [4] are better than other color spaces for image recognition. However, there is no common opinion about which of the four is consistently the best choice for color face recognition. An experiment on AR database is conducted to study these 4 popular color spaces using ERE-based CCF method. As shown in Fig. 4, HSV color space outperforms the others. Thus, our experiments make use of HSV color space. Using the other three color spaces, the relative performances of different color fusion approaches are similar to those using HSV space.

B. Face Recognition Under Different Variations

To validate the effectiveness of the proposed CCF approach for face recognition tasks under illumination variation, pose variation and mixed variations, we conduct 3 experiments on MultiPIE and AR databases.

Illumination: 18 flash-only (illumination 1-18) frontal images with neutral expression per subject in the Multi-PIE database are used in this experiment, which produces $18 \times 4 \times 105 = 7560$ images for training and testing. According to [30], if a class does not have sufficient training samples to represent some variations of its query image, they are represented by the non-class-specific component of other classes. To fully represent all variations for all classes in the training step, 4 different illuminations are randomly selected from 18 illuminations per

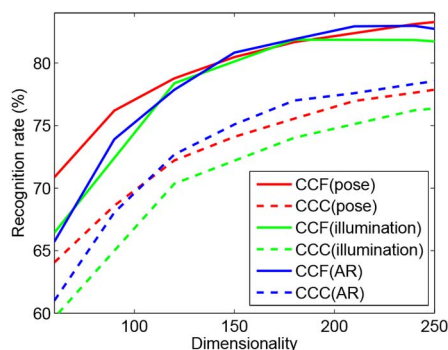


Fig. 5. FR rates of PCA-based methods using image-pixel values.

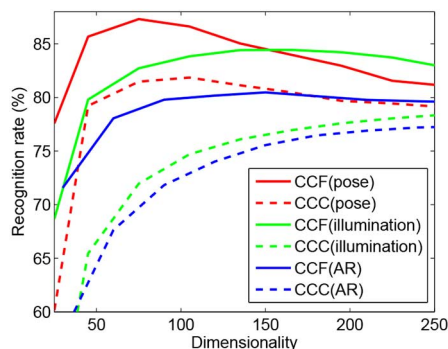


Fig. 6. FR rates of ERE-based methods using image-pixel values.

subject in session 1. Images in the remaining 3 sessions are used for testing.

Pose: For each subject in the pose subset of CMU Multi-PIE, 20 images of neutral expression are captured by 5 cameras (from -30% to $+30\%$) and flashed by the capturing camera across 4 sessions. To do training, 3 different poses are randomly selected from 5 poses per subject in session 1. Images in the remaining 3 sessions are used for testing.

Mixed Variations: 1400 undisguised images from the AR database are used in this experiment. 3 different variations are randomly selected from 7 mixed variations per subject in session 1 for training. Images in session 2 are used for testing.

Results and Analysis: The face recognition rates are averaged over 10 rounds of random selection of the training samples. The best recognition rate among all dimensions and its dimensionality at which the best recognition rate is achieved are shown in Table I for image-pixel value and in Table II for CLGWs feature. They show that the proposed CCF approach outperforms CCC and WDF methods consistently for all image variations, dimension reduction techniques and features. What's more, the dimensionality of the best recognition rate obtained by CCF is much lower than that of CCC. Although WDF gets its peak recognition rate at lower dimensionality in some cases, its recognition rate is always the lowest among the three methods. To provide more details, we plot the recognition rates against the dimensionality in Fig. 5 and 6 for color feature fusion methods using image-pixel values and in Fig. 7 and 8 for methods using CLGWs. The proposed CCF method outperforms the CCC method consistently over all dimensionality. The performance gains are significant for small number of features.

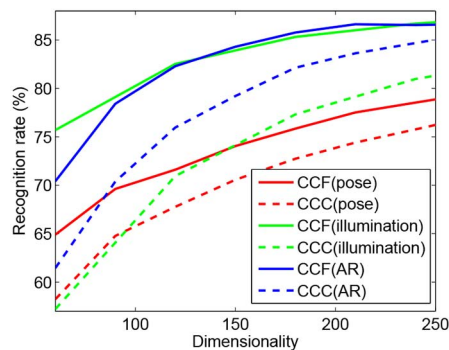


Fig. 7. FR rates of PCA-based methods using CLGWs.

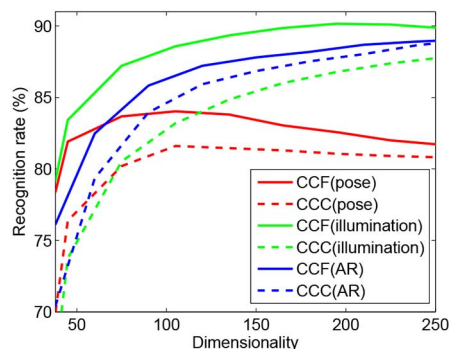


Fig. 8. FR rates of ERE-based methods using CLGWs.

TABLE I
BEST RECOGNITION RATE (%) & ITS DIMENSION USING PIXEL VALUE

Database	illumination		pose		AR	
	PCA	ERE	PCA	ERE	PCA	ERE
CCF	81.85	84.43	83.64	87.31	82.97	80.47
(dimensionality)	(180)	(165)	(300)	(75)	(240)	(150)
CCC	77.71	78.81	79.35	81.85	79.31	77.31
(dimensionality)	(360)	(300)	(360)	(105)	(390)	(270)
WDF	69.22	75.49	76.37	83.69	77.19	68.50
(dimensionality)	(120)	(300)	(210)	(45)	(240)	(180)

TABLE II
BEST RECOGNITION RATE (%) & ITS DIMENSION USING CLGWs

Database	illumination		pose		AR	
	PCA	ERE	PCA	ERE	PCA	ERE
CCF	87.61	90.16	80.90	84.03	86.63	89.39
(dimensionality)	(420)	(195)	(360)	(105)	(270)	(285)
CCC	86.05	88.24	79.14	81.60	86.37	89.11
(dimensionality)	(600)	(300)	(420)	(105)	(360)	(285)
WDF	80.40	86.03	75.03	78.27	78.09	83.20
(dimensionality)	(360)	(300)	(330)	(300)	(240)	(285)

IV. CONCLUSION

In this paper, a color channel fusion method is proposed to make use of reliability and importance of features in different color channels. By integrating the dimension reduction rule of a single color channel across all three color channels, a more effective channel fusion method is achieved. Extensive experiments on three color face datasets are conducted to validate the effectiveness and robustness of the proposed CCF method. It outperforms the CCC method and the WDF method consistently for two different dimension reduction approaches, two different types of features and 3 image variations: illumination, pose and mixed variations.

REFERENCES

- [1] L. Torres, J. Y. Reutter, and L. Lorente, "The importance of the color information in face recognition," in *Proc. Int. Conf. Image Process.*, 1999, vol. 3, pp. 627–631.
- [2] M. Rajapakse and J. Tan, "Color channel encoding with NMF for face recognition," in *Proc. Int. Conf. Image Process.*, 2004, pp. 2007–2010.
- [3] J. Choi, Y. Ro, and K. Plataniotis, "Color face recognition for degraded face images," *IEEE Trans. Syst., Man, Cybern. B: Cybern.*, vol. 39, no. 5, pp. 1217–1230, Oct. 2009.
- [4] H. Demirel and G. Anbarjafari, "Pose invariant face recognition using probability distribution functions in different color channels," *Signal Process. Lett.*, pp. 537–540, 2008.
- [5] J. Y. Choi, Y. M. Ro, and K. Plataniotis, "Color local texture features for color face recognition," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1366–1380, Mar. 2012.
- [6] C. Liu, "Extracting discriminative color features for face recognition," *Pattern Recogn. Lett.*, pp. 1796–1804, Oct. 2011.
- [7] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognit. Neurosci.*, vol. 3, no. 1, pp. 71–86, Jan. 1991.
- [8] X. D. Jiang, B. Mandal, and A. Kot, "Eigenfeature regularization and extraction in face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 3, pp. 383–394, Mar. 2008.
- [9] J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Boosting color feature selection for color face recognition," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1425–1434, May 2011.
- [10] S. H. Lee, J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Local color vector binary patterns from multichannel face images for face recognition," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2347–2353, Apr. 2012.
- [11] X. D. Jiang, "Linear subspace learning-based dimensionality reduction," *IEEE Signal Process. Mag.*, vol. 28, no. 2, pp. 16–26, Feb. 2011.
- [12] 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.
- [13] S. H. Lee, J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Local color vector binary patterns from multichannel face images for face recognition," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2347–2353, Apr. 2012.
- [14] J. Y. Choi, W. D. Neve, K. N. Plataniotis, and Y. M. Ro, "Collaborative face recognition for improved face annotation in personal photo collections shared on online social networks," *IEEE Trans. Multimedia*, vol. 31, no. 1, pp. 14–28, Feb. 2011.
- [15] J. R. Smith, "Integrated spatial and feature image systems: Retrieval, compression and analysis," Ph.D. dissertation, Columbia Univ., New York, NY, USA, 1997.
- [16] O. Ikeda, "Segmentation of faces in video footage using HSV color for face detection and image retrieval," in *Proc. Int. Conf. Image Process.*, 2003, vol. 3, pp. 913–916.
- [17] G. Paschos, "Perceptually uniform color spaces for color texture analysis: An empirical evaluation," *IEEE Trans. Image Process.*, vol. 10, no. 6, pp. 932–937, Jun. 2001.
- [18] X. D. Jiang, "Asymmetric principal component and discriminant analyses for pattern classification," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 31, no. 5, pp. 931–937, May 2009.
- [19] J. H. Friedman, "Regularized discriminant analysis," *J. Amer. Statist. Assoc.*, vol. 84, no. 405, pp. 165–175, Mar. 1989.
- [20] X. D. Jiang, B. Mandal, and A. Kot, "Enhanced maximum likelihood face recognition," *Electron. Lett.*, vol. 42, no. 19, pp. 1089–1090, 2006.
- [21] X. D. Jiang, B. Mandal, and A. C. Kot, "Complete discriminant evaluation and feature extraction in kernel space for face recognition," *Mach. Vis. Applicat.*, vol. 20, no. 1, pp. 35–46, 2009.
- [22] A. Martinez and R. Benavente, "The AR face database," CVC Tech. Rep., no. 24, Jun. 1998.
- [23] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multiple," *Image Vis. Comput.*, vol. 28, no. 5, pp. 807–813, 2010.
- [24] Z. Liu and C. Liu, "Fusion of color, local spatial and global frequency information for face recognition," *Pattern Recogn.*, pp. 2882–2890, 2010.
- [25] J. M. Chaves-González, M. A. Vega-Rodríguez, J. A. Gómez-Pulido, and J. M. Sánchez-Pérez, "Detecting skin in face recognition systems: A colour spaces study," *Digital Signal Process.*, pp. 806–823, 2010.
- [26] I. Osamu, "Segmentation of faces in video footage using HSV color for face detection and image retrieval," in *Proc. Int. Conf. Image Process.*, 2003, vol. 3, pp. 913–916.
- [27] G. Paschos, "Perceptually uniform color spaces for color texture analysis: An empirical evaluation," *IEEE Trans. Image Process.*, vol. 10, no. 6, pp. 932–937, Jun. 2001.
- [28] J. Yang, C. Liu, and L. Zhang, "Color space normalization: Enhancing the discriminating power of color spaces for face recognition," *Pattern Recogn.*, pp. 1454–1466, 2010.
- [29] L. El. Shafey, C. McCool, R. Wallace, and S. Marcel, "A scalable formulation of probabilistic linear discriminant analysis: Applied to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 7, pp. 1788–1794, Jul. 2013.
- [30] X. D. Jiang and J. Lai, "Sparse and dense hybrid representation via dictionary decomposition for face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 5, pp. 1067–1079, May 2015.