AN EFFECTIVE COLOR SPACE FOR FACE RECOGNITION

Ze Lu, Xudong Jiang and Alex Kot

School of Electrical and Electronic Engineering
Nanyang Technological University
639798 Singapore

ABSTRACT

The three color components specifying a color can be defined in various ways leading to significantly different classification abilities. Several effective color spaces including RQCr, DCS and ZRG have been proposed to achieve better face recognition performance. However, their performance is not consistent on different databases. What's more, the framework of effective color spaces has not been thoroughly studied yet. In this paper, we propose an effective color space LC_1C_2 based on a framework of effective color spaces. LC_1C_2 consists of one discriminant luminance component L and two discriminant chrominance components C_1C_2. To find the discriminant luminance component, 4 luminance components from existing effective color models are compared. After that, the weighted color space normalization technique (WCSN) is applied on the DCS color space to generate two complementary and discriminative chrominance components. Experiments conducted on three databases (FRGC, AR and CMU Multi-PIE) show that the proposed color space LC_1C_2 achieves the best face recognition performance consistently.

***Index Terms*—** Color space, face recognition

1. INTRODUCTION

Color information plays a discriminative and complementary role in face recognition process. Torres et al. [1] applied a modified PCA scheme to face recognition and their results show that the use of color information improves the recognition rate compared to the same scheme using the luminance information only. The improvement can be significant when large facial expression and illumination variations are present or the resolution of face images is low [2, 3]. Since then, considerable research efforts have been devoted to the efficient utilization of facial color information to improve the recognition performance [4].

The color space uniquely specifying a color is defined by a combination of 3 color components. Different color spaces possess significantly different characteristics and effectiveness in terms of discriminating power for visual classification tasks [5]. Various color spaces have been proposed to find the optimal way of representing color images for face recognition. In the early studies, color configurations were made through a combination of intuition and empirical comparisons without any systematic selection strategy, such as YUV [1], YCbCr, YIQ [6] and the hybrid color space YQCr in [7]. Among all possible color-component configurations discussed in [3], RQCr was proven to show the best face recognition performance. In order to derive an effective color image representation based on theoretical and experimental justifications, authors in [8] sought 3 sets of optimal coefficients to combine the R, G and B components based on a discriminant criterion and proposed the Discriminant Color Space (DCS). In addition, they also proposed two CSN (color space normalization) techniques and the ZRG color space which achieves the best FR (face recognition) performance of all normalized color spaces evaluated in [9]. RQCr and ZRG are also considered to be two most effective color representations devised for the purpose of FR in [10].

Above mentioned color spaces do achieve better face recognition performance on some databases than the others. However, the performance is not consistent on different databases. Besides, RQCr was proposed without any selection strategy or learning process, while learning-based methods ZRG and DCS follow totally different criterias. Can we find a framework to produce an effective color space which outperforms the others consistently on some popular databases? By analysing effective color spaces, we notice that they are all composed of one discriminative luminance component and two complementary chrominance components. This configuration reduces the correlation of the three color components and thus enhances the discriminating power of color spaces [9]. Based on this framework, an effective color space LC_1C_2 is proposed in this paper. Four luminance components are firstly compared to choose the most discriminant one as L. Then, our proposed weighted CSN (WCSN) technique is applied to the DCS [8] color space to derive two discriminative and complementary chrominance components C_1C_2.

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, Singapore, under its Interactive Digital Media (IDM) Strategic Research Programme.

2. THE PROPOSED COLOR SPACE LC_1C_2

In this section, a framework of effective color spaces for face recognition tasks is discussed first and then an effective color space LC_1C_2 is proposed.

2.1. A framework of effective Color Spaces for Face Recognition Tasks

Rather than merely searching for a more effective color space as the previous research [1, 3, 7, 8, 9], we try to explore the framework of effective color spaces for face recognition tasks. According to previous studies, $I_1I_2I_3$ [11], YUV, YIQ, YCbCr, [1, 12], LSLM [13], RQCr [3], YQCr [7] and ZRG [9] are relatively effective color spaces. By analysing the 3 components of them, we find that they are all composed of one luminance component (I_1, Y, L, R) and two chrominance components (I_2I_3, UV, IQ, CbCr, SLM, QCr, ZG). It is explained in [14] that the luminance structure of face images is of great significance, color cues play a complementary role in face recognition and their contribution becomes evident when shape cues are degraded. Thus when we combine luminance and chrominance components, the recognition performance is significantly better than that with grayscale images only. Further validations of this configuration could be found in [9] that the separation of luminance and chrominance information reduces the correlation between the three color components and makes the discriminative information contained in the three color components as mutually complementary as possible. Thus, the concatenation of three components effectively makes use of the discriminative information from three color channels.

Let L, C_1, C_2 be the three color components derived by the linear transformation of the RGB color space, the framework of constructing a effective color space for face recognition is to select a discriminative luminance component L, and two discriminative chrominance components C_1, C_2 to reduce the correlation between L and C_1, C_2.

$$\begin{bmatrix} L(luminance) \\ C_1(chrominance) \\ C_2(chrominance) \end{bmatrix}$$

2.2. A Discriminative Luminance Component L

In many face recognition algorithms, a color image in the RGB color space is converted into a monochrome image by linearly combining its three color components [15]. However, theoretical and experimental justifications are lacking for investigating which monochrome image is the best representation of the color image for the recognition purpose. In this section, we compare 4 effective and commonly used luminance components including I_1 from $I_1I_2I_3$ [16], R from RGB, Y from YUV [12] and L from LSLM [13]. They are calculated from the RGB color space in (1)-(4).

The $I_1I_2I_3$ color space proposed by Ohta et al. applies a K-L transformation to decorrelate the RGB components. The effectiveness of the luminance component I_1 for face verification is also shown in [11]. But I_1 implicitly assumes a uniform distribution of color values over the entire color space. For a task such as face recognition, color values tend to be more tightly confined to a small portion of the color space [17].

$$I_1 = [1/3, 1/3, 1/3][R, G, B]^t \tag{1}$$

The component R in the RGB color space has been shown to be more effective than other color components such as I_1 and Gray for face recognition [6]. In addition, the R channel of skin-tone color is known to be the best monochrome channel for face recognition [18]. As R discards useful information by ignoring G and B components, the superiority of R component over the others is only reflected on the FRGC database [19].

$$R = [1, 0, 0][R, G, B]^t \tag{2}$$

YUV, YIQ, YCbCr are three color standards commonly used for video transmission efficiency. In these 3 color spaces, the RGB components are separated into luminance component Y and remaining chrominance components. Y performs the best for use to display pictures on monochrome (black and white) televisions [20]. However, it may not be optimal to be utilized in face recognition. The LSLM color space is a linear transformation of the RGB color space based on the opponent signals of the cones: blackwhite, redgreen, and yellowblue . L describes the luminance information.

$$Y = [0.299, 0.587, 0.114][R, G, B]^t \tag{3}$$

$$L = [0.209, 0.715, 0.076][R, G, B]^t \tag{4}$$

According to the normalized response spectra of human cone cells in [21], the eye is most sensitive to green light than other colors because this stimulates the two most common (M and L) of the three kinds of cones at all light levels. That's also the reason why most of the sensors in cameras are green sensors. Therefore when we linearly combine R, G, B components, assigning more weight to the G component is expected to produce a more discriminative luminance component. This provides a justification that L should be adopted as the luminance component rather than Y.

In order to further validate the effectiveness of the L component experimentally, we conduct the discriminating power analysis on I_1, R, Y and L components. Similar to [9], the discriminating power of color images corresponding to a given color component is characterized by the discriminant criterion value J defined in (5), (6) and (7). Suppose a face image is rearranged into a row vector I, then I_{ij} is the jth image vector in class i, where $i = 1, 2, ..., p$ and $j = 1, 2, ..., q$. $\overline{I_i}$

Table 1. Discriminant value of I_1, R, Y and L evaluated on 13990 face images.

Component	I_1	R	Y	L
J	0.2943	0.2039	0.2928	0.3043

is the mean vector in class i and $\overline{I}$ is the mean vector of all training samples.

$$V_b = \frac{1}{p}\sum_{i=1}^{p}(\overline{I_i} - \overline{I})(\overline{I_i} - \overline{I})^t. \tag{5}$$

$$V_w = \frac{1}{pq}\sum_{i=1}^{p}\sum_{j=1}^{q}(I_{ij} - \overline{I_i})(I_{ij} - \overline{I_i})^t. \tag{6}$$

$$J = \frac{V_b}{V_w}. \tag{7}$$

Using (5), (6) and (7), the discriminant criterion value J of I_1, R, Y and L components shown on Table. 1 are derived. We collect 13990 training images from the training part of FRGC, AR [22] and CMU Multi-PIE [23] databases. Table. 1 verifies that the discriminating power of the L component is greater than that of I_1, R or Y component.

2.3. Two Discriminative Chrominance Components C_1C_2

Suppose M_1, M_2, M_3 are three color components derived by the following linear transformation of the RGB color space in (8):

$$\begin{bmatrix} M_1 \\ M_2 \\ M_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{8}$$

Chrominance is the signal used in video systems to convey the color information of the picture, separately from the accompanying luminance signal [24]. It is usually represented as two color-difference components such as $U = B - Y$ (blue - luma) and $V = R - Y$ (red - luma) from the YUV color space. In other words, color components such as B or R contains the same luminance information Y. Thus when $s_i = \sum_{j=1}^{3} a_{ij} = 0, (i = 1, 2, 3)$, the luminance information Y would get cancelled and the remaining information forms the chrominance components U and V.

In order to generate such two components C_1C_2, the across-color-component normalization technique (CSN-II) [9] shown in algorithm 1 is adopted. The CSN-II technique normalizes any color space transformation matrix $[A_1, A_2, A_3]^t$, so that for the two normalized color space transformation vectors B_i, $sb_i = \sum_{j=1}^{3} b_{ij} = 0, (i = 1, 2)$. However, there exist two problems of this technique. The first problem is how to choose the initial color components or how to choose A_i or a_{ij}. In [9], initial color components that determine A_i or a_{ij} are sought from RGB and XYZ to find the best normalized color space ZRG. Is RGB or XYZ the only choice for the initial color space? Secondly, in the 3rd and 4th steps of algorithm 1, since different initial color components M_i possess different discriminating power, A_i should be weighted accordingly.

Algorithm 1 CSN-II

1: $s_i = \sum_{j=1}^{3} a_{ij}, i = 1, 2, 3$
2: $[s_1, s_2, s_3][k_1, k_2, k_3]^t = 0$, the solution has two basis vectors $K_1 = [k_{11}, k_{12}, k_{13}], K_2 = [k_{21}, k_{22}, k_{23}]$
3: $B_1 = [b_{11}, b_{12}, b_{13}] = \sum_{i=1}^{3} k_{1i}A_i$
4: $B_2 = [b_{21}, b_{22}, b_{23}] = \sum_{i=1}^{3} k_{2i}A_i$

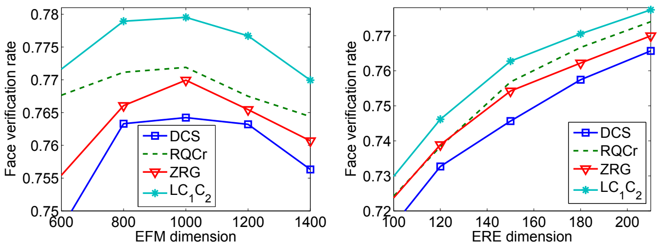
The first problem of the CSN-II technique comes before the color space normalization, different color components from RGB and XYZ color spaces were combined to find the color configuration which generates the best normalized color space, ZRG, in [9]. However, there is no reason why RGB and XYZ color spaces should be selected. Since the purpose of this section is to derive two discriminative chrominance components, the discriminative information in the initial color components should be effectively fused. To maximize the difference of discriminating power possessed by 3 initial color components, the DCS color space [8] is employed in our framework to determine A_i or a_{ij} in algorithm 1. As for the other problem, when A_i are linearly combined in the 3rd and 4th steps of algorithm 1, weights should be assigned to A_i according to their discrimination power. That is exactly what the eigenvalue indicates in the eigen-decomposition step of LDA in the DCS color space [8]. So we multiply A_i with their corresponding eigenvalue λ_i to effectively utilize the discriminative information. Therefore, the 3rd and 4th steps of algorithm 1 become (9) and (10):

$$B_1 = [b_{11}, b_{12}, b_{13}] = \sum_{i=1}^{3} k_{1i}\lambda_i A_i \tag{9}$$

$$B_2 = [b_{21}, b_{22}, b_{23}] = \sum_{i=1}^{3} k_{2i}\lambda_i A_i \tag{10}$$

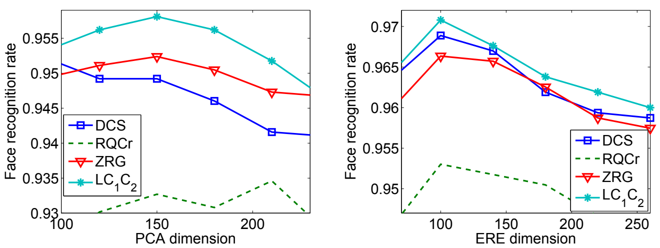
3. EXPERIMENTS

The effectiveness of our proposed color space LC_1C_2 is assessed on three databases, FRGC, CMU Multi-PIE and AR. It is compared with other 3 state-of-art color spaces including DCS, RQCr and ZRG. Three color components are concatenated into one pattern vector to combine the information in them. A basic image normalization method by removing the mean and normalizing the standard deviation of each component is used before the concatenation to avoid the negative effect of magnitude dominance of one component over the others. Both EFM [25] and ERE [26] are adopted as the dimension reduction methods on FRGC because all other color



(a) (b)

Fig. 1. Face verification rates against (a) EFM dimension and (b) ERE dimension on FRGC 2.0 database.



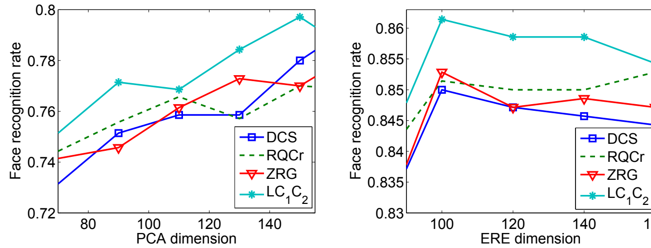
(a) (b)

Fig. 2. Face recognition rates against (a) PCA dimension and (b) ERE dimension on pose variation subset of CMU Multi-PIE database.

FR methods make use of the EFM method for extracting low-dimensional features on FRGC[3, 8, 9] and ERE outperforms all other FR methods discussed in [10, 26]. On CMU Multi-PIE and AR databases, PCA and ERE are used for dimension reduction and the minimum-Mahalanobis-distance classifier is used for classifying all query images. It is shown in [27, 28] that PCA may significantly enhance the recognition accuracy.

There are 12,776 training images, 16,028 controlled target images, and 8,014 uncontrolled query images in FRGC Experiment 4. The controlled images have good image quality, while the uncontrolled images display poor image quality, such as large illumination variations, low resolution, and blurring. The face recognition performance is reported by the face verification rates at FAR=0.1% according to recent published studies. In addition, face images are first cropped from the original images and resized to a spatial resolution of 32×32 to be consistent with [9]. Experimental results shown on Fig. 1(a) and Fig. 1(b) indicate that our proposed LC_1C_2 color space outperforms DCS, RQCr and ZRG color spaces consistently over all EFM and ERE feature dimensions.

For the pose variation subset of CMU Multi-PIE, face images are captured across 4 sessions. The first 105 subjects which appear in all 4 sessions are used in the experiments. For each subject, 20 images of neutral expression are captured by 5 cameras (from -30% to +30%). Images are cropped based



(a) (b)

Fig. 3. Face recognition rates against (a) PCA dimension and (b) ERE dimension on AR database.

on the eye locations and resized to 50×40 similar to [29]. All images in session 1 are used to do training and images in the remaining 3 sessions are used for testing. Experimental results are shown in Fig. 2(a) and Fig. 2(b). Similarly, the proposed LC_1C_2 color space surpasses DCS, RQCr and ZRG color spaces consistently against all PCA and ERE feature dimensions.

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 in our experiments, so there are in total 1400 undisguised images from the AR database. The face region is cropped manually from original images and resized to 55×40 [29]. Also, 4 different variations are randomly selected from 7 mixed variations per subject in session 1 for training. Images in session 2 are used for testing. From the experimental results shown in Fig. 3(a) and Fig. 3(b), we can see that the proposed LC_1C_2 color space once again performs better than other color spaces consistently against the PCA and ERE dimension.

4. CONCLUSION

In this paper, we first propose a framework of effective color spaces for face recognition, which consists of one luminance component and two chrominance components. Based on the framework, we propose an effective color space LC_1C_2. According to the experimental results on FRGC, CMU Multi-PIE and AR databases, the proposed color space LC_1C_2 outperforms DCS, RQCr and ZRG color spaces consistently using different dimension reduction methods on 3 databases.

5. REFERENCES

[1] L. Torres, J. Reutter, and L. Lorente, "The importance of the color information in face recognition," in *IEEE Intl. Conf. Image Processing*, 1999, vol. 3, pp. 627–631.

[2] M. Rajapakse, J. Tan, and J. Rajapakse, "Color channel

encoding with nmf for face recognition," in *IEEE Intl. Conf. Image Processing*, 2004, vol. 3, pp. 2007–2010.

[3] J. Y. Choi, Y. M. Ro, and K. N Plataniotis, "Color face recognition for degraded face images," *IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 39, no. 5, pp. 1217–1230, 2009.

[4] Z. Lu, X. Jiang, and A. C Kot, "A color channel fusion approach for face recognition," *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 1839–1843, 2015.

[5] R. Lukac and K. N Plataniotis, *Color image processing: methods and applications*, CRC press, 2006.

[6] P. Shih and C. Liu, "Comparative assessment of content-based face image retrieval in different color spaces," *Intl. Journal of Pattern Recognition and Artificial Intelligence*, vol. 19, no. 07, pp. 873–893, 2005.

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

[8] J. Yang and C. Liu, "A discriminant color space method for face representation and verification on a large-scale database," in *IEEE Intl. Conf. Pattern Recognition*, 2008, pp. 1–4.

[9] J. Yang, C. Liu, and L. Zhang, "Color space normalization: Enhancing the discriminating power of color spaces for face recognition," *Pattern Recognition*, vol. 43, no. 4, pp. 1454–1466, 2010.

[10] J. Y. Choi, Y. M. Ro, and K. N Plataniotis, "Color local texture features for color face recognition," *IEEE Trans. Image Processing*, vol. 21, no. 3, pp. 1366–1380, 2012.

[11] J. Kittler and M. T Sadeghi, "Physics-based decorrelation of image data for decision level fusion in face verification," in *Multiple Classifier Systems*, pp. 354–363. Springer, 2004.

[12] W. H Buchsbaum, *Color TV servicing*, Prentice Hall, 1975.

[13] P. Colantoni et al., "Color space transformations," *see http://www.radugaryazan.ru/files/doc /colorspacetransform95.pdf*, 2004.

[14] A. Yip and P. Sinha, "Role of color in face recognition," 2001.

[15] J. Yang and C. Liu, "A general discriminant model for color face recognition," in *IEEE Intl. Conf. Computer Vision*, 2007, pp. 1–6.

[16] Y. Ohta, *Knowledge-based interpretation of outdoor natural color scenes*, vol. 4, Morgan Kaufmann, 1985.

[17] C. F Jones III and A L. Abbott, "Optimization of color conversion for face recognition," *EURASIP Journal on Applied Signal Processing*, vol. 2004, pp. 522–529, 2004.

[18] M. T Sadeghi, S. Khoshrou, and J. Kittler, "Confidence based gating of colour features for face authentication," in *Multiple Classifier Systems*, pp. 121–130. Springer, 2007.

[19] P J. Phillips, P. J Flynn, T. Scruggs, K. W Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *IEEE Conf. Computer Vision and Pattern Recognition*, 2005, vol. 1, pp. 947–954.

[20] J. Lu and K. N Plataniotis, "On conversion from color to gray-scale images for face detection," in *IEEE Conf. Computer Vision and Pattern Recognition Workshops*, 2009, pp. 114–119.

[21] G. Wyszecki, VS Stiles, and K. L Kelly, "Color science: concepts and methods, quantitative data and formulas," *Physics Today*, vol. 21, no. 6, pp. 83–84, 2009.

[22] A. M Martinez, "The ar face database," *CVC Technical Report*, vol. 24, 1998.

[23] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-pie," *Image and Vision Computing*, vol. 28, no. 5, pp. 807–813, 2010.

[24] V. Georges, "System of television in colors," May 15 1945, US Patent 2,375,966.

[25] C. Liu and H. Wechsler, "Robust coding schemes for indexing and retrieval from large face databases," *IEEE Trans. Image Processing*, vol. 9, no. 1, pp. 132–137, 2000.

[26] X. Jiang, B. Mandal, and A. Kot, "Eigenfeature regularization and extraction in face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 30, no. 3, pp. 383–394, 2008.

[27] X. Jiang, "Asymmetric principal component and discriminant analyses for pattern classification," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, no. 5, pp. 931–937, 2009.

[28] X. Jiang, "Linear subspace learning-based dimensionality reduction," *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 16–26, 2011.

[29] Xudong Jiang and Jian Lai, "Sparse and dense hybrid representation via dictionary decomposition for face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 37, no. 5, pp. 1067–1079, 2015.