HOLISTIC ORTHOGONAL ANALYSIS OF DISCRIMINANT TRANSFORMS FOR COLOR FACE RECOGNITION

Xiaoyuan Jing[1,2], Qian Liu[1], Chao Lan[1], Jiangyue Man[1], Sheng Li[1], David Zhang[3]
[1]School of Automation, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China
[2]State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, 210093, China
[3]Department of Computing, Hong Kong Polytechnic University, Kowloon, Hong Kong

ABSTRACT

The key of color face recognition technique is how to effectively utilize the complementary information between color components and remove their redundancy. Present color face recognition methods generally reduce the correlations between color components in the image pixel level, and then extract the discriminant features from the uncorrelated color face images. In this paper, we propose a novel color face recognition approach based on the holistic orthogonal analysis (HOA) of discriminant transforms of color images. HOA can reduce the correlation of color information in the feature level. It in turn achieves the discriminant transforms of red, green and blue color images by using the Fisher criterion, and simultaneously makes the achieved transforms mutually orthogonal. Experimental results on the AR and FRGC-2 public color face image databases demonstrate that the proposed approach acquires better recognition performance than several representative color face recognition methods.

Index Terms— Color face recognition, Discriminant transforms, Holistic orthogonal analysis, Feature extraction.

1. INTRODUCTION

Color images are increasingly used in the fields of face recognition [1]. The key of color face recognition technique is how to effectively utilize the complementary information between color images and remove their redundancy. Yang and Liu [2] presented an extended general color image discriminant (Extended GCID) algorithm that produces three groups of weights to fuse color images and then extracts discriminant features from the fused images. Liu [3] presented the uncorrelated color space (UCS), the independent color space (ICS), and the discriminating color space (DCS) methods for face recognition. UCS applies the principal component analysis (PCA) to reduce the correlations between color components. ICS assumes that each color image is defined by three independent source images that can be derived by a blind source separation procedure, such as the independent component analysis (ICA). And DCS applies the discriminant analysis [4] to define three new component images that are effective for pattern recognition. Shih and Liu transformed the RGB color space to the YQCr color configuration and extracted discriminant features for classification [5]. Yang et al. transformed RGB space to the HSV color space, and then employed the Hue and Saturation components to do complex PCA transform [6].

Present color recognition methods generally reduce the correlations of color components or transform RGB space to other spaces in the image pixel level, and then extract discriminant features. In this paper, we propose a novel color face recognition approach which reduces the correlation of color information in the feature level. It is based on the holistic orthogonal analysis (HOA) of discriminant transforms of color images. It in turn acquires the discriminant transforms of red, green and blue color images by using the Fisher criterion, and simultaneously makes the achieved transforms mutually orthogonal. Experiments on the AR [7] and FRGC-2 [8] public color face databases validate the effectiveness of the proposed approach.

2. HOLISTIC ORTHOGONAL ANALYSIS OF DISCRIMINANT TRANSFORMS (HOA)

We first use HOA to acquire three discriminant transforms of red, green and blue components, respectively. Then we describe the realization algorithm of HOA and related theoretical property.

2.1. Discriminant transform of red component

Based on the Fisher criterion, we calculate the discriminant transform of red component W_R by

$$\max\ J(W_R) = \frac{\left|W_R^T S_{bR} W_R\right|}{\left|W_R^T S_{wR} W_R\right|}, \tag{1}$$

where $|\bullet|$ expresses the determinant of a square matrix, S_{bR} is the between-class scatter matrix of red component, and S_{wR} is the within-class scatter matrix of red component.

Therefore we can achieve W_R by solving the following eigenequation:

$$P_R W_R = \lambda W_R,\ P_R = S_{wR}^{-1} S_{bR}. \tag{2}$$

Hence W_R is a matrix that consists of d_R eigenvectors corresponding to d_R different nonzero eigenvalues of P_R.

2.2. Discriminant transform of green component

For eliminating the correlation of discriminant transforms of red and green components, we consider $W_G^T W_R = 0$, where W_G is the discriminant transform of green component, being a constraint on the Fisher criterion:

$$\max \ J(W_G) = \frac{\left|W_G^T S_{bG} W_G\right|}{\left|W_G^T S_{wG} W_G\right|}, \quad s.t. \quad W_G^T W_R = 0 \qquad (3)$$

where S_{bG} is the between-class scatter matrix of green component, and S_{wG} is the within-class scatter matrix of green component.

Firstly, we put forward a generalized theorem to solve the following objective function:

Theorem 1:

$$\max \ J(W_2) = \frac{\left|W_2^T S_b W_2\right|}{\left|W_2^T S_w W_2\right|}, \quad s.t. \quad W_2^T W_1 = 0 \qquad (4)$$

where W_1, S_b and S_w are all known matrices. W_2 is achieved by solving the following eigenequation:

$$PW_2 = \lambda W_2, \ P = S_w^{-1}\left(I - W_1(W_1^T S_w^{-1} W_1)^{-1} W_1^T S_w^{-1}\right) S_b, \quad (5)$$

where I is a unit matrix. Thus W_2 is a matrix that consists of d eigenvectors corresponding to d different nonzero eigenvalues of P.

The proof is given in Appendix A.

Secondly, Formula (4) can be changed into Formula (3) via separately replacing W_1, W_2, S_b and S_w by W_R, W_G, S_{bG} and S_{wG}. Thus we obtain Theorem 2 to solve W_G:

Theorem 2:

W_G is achieved by solving the following eigenequation:

$$P_G W_G = \lambda W_G, \ P_G = S_{wG}^{-1}\left(I - W_R(W_R^T S_{wG}^{-1} W_R)^{-1} W_R^T S_{wG}^{-1}\right) S_{bG}, \quad (6)$$

where I is a unit matrix. W_G consists of d_G eigenvectors corresponding to d_G different nonzero eigenvalues of P_G.

The proof is similar to that of Theorem 1.

2.3. Discriminant transform of blue component

For eliminating the correlation of between blue discriminant transform and foregoing two discriminant transforms, we construct the following objective function and constraints:

$$\max \ J(W_B) = \frac{\left|W_B^T S_{bB} W_B\right|}{\left|W_B^T S_{wB} W_B\right|}, \quad s.t. \quad W_B^T W_R = 0, \ W_B^T W_G = 0 \qquad (7)$$

where W_B is the discriminant transform of blue component, S_{bB} is the between-class scatter matrix of blue component, and S_{wB} is the within-class scatter matrix of blue component.

Formula (4) can be transformed into Formula (7) via separately replacing W_1, W_2, S_b and S_w by $W = [W_R, \ W_G]$, W_B, S_{bB} and S_{wB}. Thus we obtain Theorem 3:

Theorem 3:

W_B is achieved by solving the following eigenequation:

$$P_B W_B = \lambda W_B, \ P_B = S_{wB}^{-1}\left(I - W(W^T S_{wB}^{-1} W)^{-1} W^T S_{wB}^{-1}\right) S_{bB}, \quad (8)$$

where $W = [W_R, \ W_G]$, I is a unit matrix. W_B consists of d_B eigenvectors corresponding to d_B different nonzero eigenvalues of P_B.

The proof is similar to that of Theorem 1.

2.4. Algorithm description

We describe the realization algorithm of HOA as follows:

Step 1: Calculate W_R by using Formula (2) on the training sample set.
Step 2: Calculate W_G by using Formula (6) on the training sample set.
Step 3: Calculate W_B by using Formula (8) on the training sample set.
Step 4: Separately orthonormalize W_R, W_G and W_B.
Step 5: Apply W_R, W_G and W_B to the red, green and blue components of all samples, respectively. For each sample, combine three extracted discriminant feature vectors into one vector.
Step 6: Use the nearest neighbor classifier with the cosine distance to do classification.

2.5. Theoretical property

Step 4 in the above algorithm description can make the discriminant vectors within the achieved discriminant transform are orthogonal. This is helpful to further reduce the correlation of discriminant features. However, can this orthogonalization operation guarantee W_G and W_B still satisfying Formulas (6) and (8)? Theorem 4 proves that orthogonalizing W_G and W_B does not influence the satisfaction of the Fisher criterion and holistic orthogonal constraints between discriminant transforms. For the simplicity, the proof of Theorem 4 takes the form of Theorem 1.

Theorem 4:

Suppose that W_1 and W_2 in Theorem 1 are orthogonalized to W_1' and W_2', respectively. We have $J(W_2') = J(W_2)$ and $W_2'^T W_1' = 0$.

The proof is given in Appendix B.

3. EXPERIMENTS

We complete the experiment on the AR public color face image database at first. This database contains 102 individuals, each 26 images. We crop every image to the size of 60×60. All cropped images of one subject are shown in Figure 1. The major differences between them are the expression, illumination, position, pose and sampling time. In order to effectively evaluate the impact of different variations to the recognition results, we in turn choose the following 2-8 representative images of every subject as training samples: (1), (2), (5), (8), (11), (14), (17) and (19), and use the remainder as testing samples.

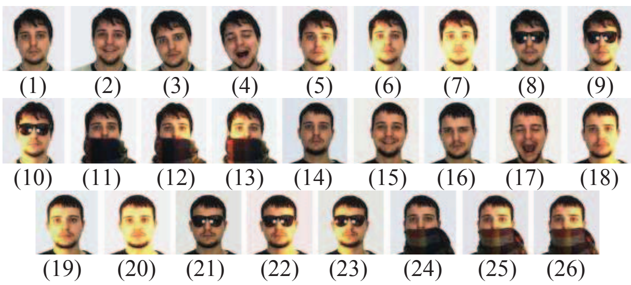

Figure 1: Demo images of one individual on AR database.

We also complete the experiment on the FRGC-2 public color face image database. This database used in the experiment contains 100 individuals, each 24 images. And we crop every sample image to the size of 60×60. Figure 2 shows all samples of one subject. We in turn choose the front 2-8 images of each person as training samples and use the remainder as testing samples.

Figure 2: Demo images of one individual on FRGC-2 database.

We compare our approach with the color Fisherface (CFF) method, UCS [4], ICS [4], DCS [4], and Extended GCID [3]. CFF is an extension of the grayscale Fisherface, which combines the red, green and blue component vectors of each sample into a vector, and then apply the Fisherface method [7] to the vector. For all compared methods, we use the nearest neighbor classifier to do classification. In the proposed approach, we employ PCA first to avoid the singular within-class scatter matrix.

Figures 3 and 4 separately show the recognition rates of the proposed HOA approach and five compared methods on the AR and FRGC-2 databases, where the number of training samples per person is ranged from 2 to 8. Table 1 shows the average recognition rates of all compared methods on the two databases.

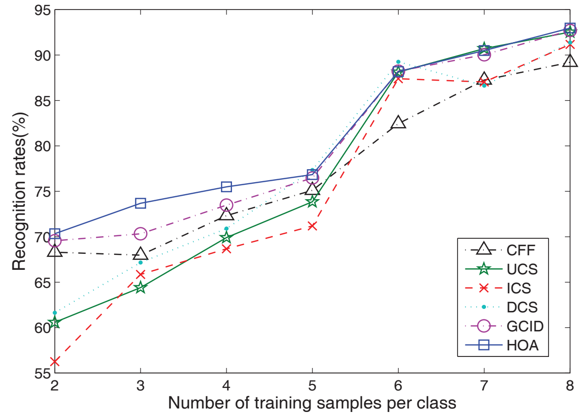

Figure 3: Recognition rates of compared methods on AR.

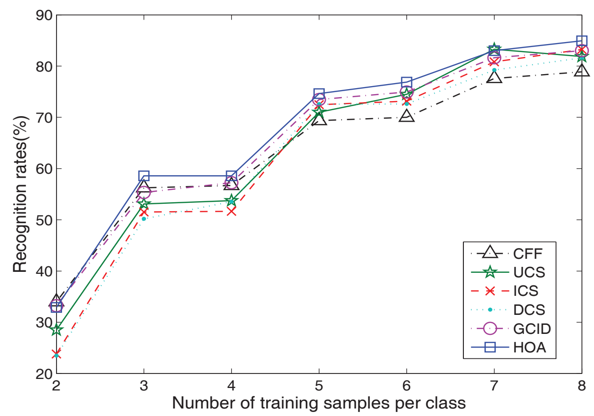

Figure 4: Recognition rates of compared methods on FRGC-2.

Table 1: Average recognition rates of all compared methods on AR and FRGC-2 databases.

Methods	Average recognition rates (%)	
	AR	FRGC-2
CFF	77.51	63.24
UCS	77.15	63.71
ICS	75.36	62.38
DCS	77.76	61.89
Extended GCID	80.11	65.55
HOA	81.04	67.06

Compared with CFF, UCS, ICS, DCS and Extended GCID, HOA separately improves the average recognition rates on both databases at least by 3.53%(=81.04%-77.51%), 3.35%(=67.06%-63.71%),4.68%(=67.06%-62.38%), 3.28% (=81.04%-77.76%), 0.93%(=81.04%-80.11%).

4. CONCLUSION

In this paper, we propose a novel color face recognition approach based on the holistic orthogonal analysis (HOA) of discriminant transforms. In the feature level, HOA makes full use of supplementary information of discriminant transforms obtained from the color images, and reduces the correlation between the transforms. Experimental results on the AR and FRGC-2 public color face image databases demonstrate that the proposed approach achieves better recognition performance than several representative color face recognition methods.

5. ACKNOWLEDGEMENTS

The work described in this paper was fully supported by the NSFC under Project No.60772059, the New Century Excellent Talents of Education Ministry under Project No. NCET-09-0162, the Doctoral Foundation of Education Ministry under Project No. 20093223110001, the Qin-Lan Engineering Academic Leader of Jiangsu Province, the Foundation of Jiangsu Province Universities under Project No.09KJB510011, the Foundation of NJUPT under Project No. NY207027, Project No.NY208051.

6. APPENDIX A

First, we construct the Lagrange function:

$$L(W_2) = W_2^T S_b W_2 - \lambda\left(W_2^T S_w W_2 - C_1\right) - \mu\left(W_2^T W_1 - C_2\right), \quad \text{(A-1)}$$

where λ and μ are the Lagrange multipliers, and C_1 and C_2 are two constant matrices.

We set the derivative of $L(W_2)$ in Eq. (A-1) on W_2 to be zero:

$$\frac{\partial L(W_2)}{\partial W_2} = 2S_b W_2 - 2\lambda S_w W_2 - \mu W_1 = 0 . \quad \text{(A-2)}$$

Multiplying Eq. (A-2) by $W_1^T S_w^{-1}$, we have

$$2W_1^T S_w^{-1} S_b W_2 - \mu W_1^T S_w^{-1} W_1 = 0 . \quad \text{(A-3)}$$

Thus μ may be expressed as

$$\mu = 2(W_1^T S_w^{-1} W_1)^{-1} W_1^T S_w^{-1} S_b W_2 . \quad \text{(A-4)}$$

Due to Eqs. (A-2) and (A-4), we have

$$S_b W_2 - \lambda S_w W_2 - W_1 (W_1^T S_w^{-1} W_1)^{-1} W_1^T S_w^{-1} S_b W_2 = 0 , \quad \text{(A-5)}$$

that is,

$$S_w^{-1}\left(I - W_1 (W_1^T S_w^{-1} W_1)^{-1} W_1^T S_w^{-1}\right) S_b W_2 = \lambda W_2 , \quad \text{(A-6)}$$

where I is a unit matrix.

Eq. (A-6) is equivalent to Formula (5). Proof is over.

7. APPENDIX B

(I) We prove $J(W_2') = J(W_2)$:

Suppose that W_2 consists of d eigenvectors corresponding to d different nonzero eigenvalues of P in Formula (5), i.e., $W_2 = [w_{21}, w_{22}, \cdots, w_{2d}]$. We orthogonalize W_2 as

$$W_2' = W_2 A^T , \quad \text{(A-7)}$$

where

$$W_2' = [w_{21}', w_{22}', \cdots, w_{2d}'] , \quad \text{(A-8)}$$

A is an orthogonalized transformation matrix whose dimension is $d \times d$. Since d eigenvectors of W_2 are linearly uncorrelated, the rank of A is d . In other words, A is a full-rank matrix.

There are several methods to realize the orthogolization. In the experiment, we employ the conventional Schmidt orthogonalization. So we have

$$\begin{aligned} J(W_2') &= \frac{\left|W_2'^T S_b W_2'\right|}{\left|W_2'^T S_w W_2'\right|} = \frac{\left|A W_2^T S_b W_2 A^T\right|}{\left|A W_2^T S_w W_2 A^T\right|} \\ &= \frac{|A|\left|W_2^T S_b W_2\right|\left|A^T\right|}{|A|\left|W_2^T S_w W_2\right|\left|A^T\right|} = \frac{\left|W_2^T S_b W_2\right|}{\left|W_2^T S_w W_2\right|} = J(W_2) \end{aligned} . \quad \text{(A-9)}$$

(II) We prove $W_2'^T W_1' = 0$:

Similar to W_2', we orthogonalize W_1 as

$$W_1' = W_1 B^T , \quad \text{(A-10)}$$

where B is an orthogonalized transformation matrix whose dimension is $d \times d$. We have

$$W_2'^T W_1' = A W_2^T W_1 B^T = A\left(W_2^T W_1\right) B^T = 0 . \quad \text{(A-11)}$$

Proof is over.

8. REFERENCES

[1] L. Torres, J. Y. Reutter, and L. Lorente, "The importance of the color information in face recognition", *Int. Conf. Image Processing,* pp. 627-631, 1999.

[2] J. Yang, and C. Liu, "Color image discriminant models and algorithms for face recognition," *IEEE Trans. Neural Networks*, 19(12): 2088-2098, 2008.

[3] C. Liu, "Learning the uncorrelated, independent, and discriminating color spaces for face recognition," *IEEE Trans. Information Forensics and Security*, 3(2): 213-222, 2008.

[4] P. N. Belhumeur, J. Hespanda, and D. Kiregeman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Trans. Pattern Analysis and Machine Intelligence,* 19(7): 711-720, 1997.

[5] P. Shih, and C. Liu, "Improving the face recognition grand challenge baseline performance using color configurations across color spaces," *Int. Conf. Image Processing*, pp. 1001-1004, 2006.

[6] J. Yang, D. Zhang, Y. Xu, and J. Yang, "Recognize color face images using complex Eigenfaces", *Int. Conf. Biometrics*, pp. 64-68, 2006.

[7] A. M. Martinez, and R. Benavente, *The AR Face Database*, Academic Press, New York, 1990.

[8] 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," *Int. Conf. Computer Vision and Pattern Recognition*, pp. 947-954, 2005.