2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)
Bari, Italy. October 6-9, 2019

Cross-Channel Similarity based Histograms of Oriented Gradients for Color Images

Zheng Zhou and Yicong Zhou
Department of Computer and Information Science
University of Macau, Macau 999078, China
Email: alexzhouzhengz@gmail.com, yicongzhou@um.edu.mo

Abstract— **The local features have gained prestige as the powerful descriptors, however, when handling color images, most existing descriptors fail to find an efficient strategy to make full use of channel correlation. To tackle this problem, we propose the cross-channel similarity based histograms of oriented gradients (CCS-HOG) model for color images. Different from the existing methods, our model integrates the color correlation with the structure features together in a better way. To find out the inner connection between channels, the cross-channel similarity measure is developed as a suitable approach. Experimental results for face recognition and kinship verification illustrate the the performance of CCS-HOG superior to other state-of-the-arts descriptors.**

I. Introduction

Extensive applications, such as, kinship verification and face recognition have attracted increasing attention in pattern recognition and machine vision [1], [2]. A key step of recognition and verification is to form the discriminative and robust descriptor, which aims to distinguish the variance among classes and explore the essential similarity within class by analyzing the images.

To extract features from the images, an enormous number of methods have been proposed, including principal component analysis (PCA) [3], sparse PCA (SPCA) [4], straightforward two-dimensional PCA (2D-PCA) [5], and linear discriminant analysis (LDA) [6]. Besides the above holistic methods based on eigenfaces and fisherfaces [6], the face recognition system using local features, such as, local binary pattern (LBP), [7] histograms of oriented gradients (HOG) [8], and gabor filters [9] also have attracted more attention due to the simplicity and the robustness to conditional variations. The HOG [8] is originally developed by Dalal and Triggs to detect pedestrians because of its resistance against orientation and illuminance variations. The main idea of HOG descriptor is to extract the structure and shape of images with the intensities of gradients and directions of edges and finally the features are represented by histograms. Recently, there has been springing out wide range of applications, such as, traffic sign detection [10], car detection [11], [12], face matching [13], face recognition [14], and activity recognition [15].

Although the HOG descriptor has achieved promising performance in applications of real scene, its performance may sharply degrade when handling color images. The main reason is that most of the descriptors are originally designed for gray-scale images. To deal with color images, one intuitive way is to regard each color channel as a gray-scale image. For example, the work in [8] performed the HOG on three color channels individually and then selected the color channel with the best performance in pedestrians detection. An obvious limitation of the above scheme is that the high correlation between three color channels are fully ignored. Another way is to process the color images independently and then concatenate the three channels together. However, it also needs to extract features from three independent color channels, which means that it still fails to make full use of the correlation between three channels. In [16], the color features are represented by the histograms of color (HoC) from the hue and saturation of color images and then fused together with other features. Therefore, how to make full use of the correlation between three channels is still a challenge.

To uncover the correlation between color channels, cross-channel similarity based histograms of oriented gradients (CCS-HOG) for color images is proposed in this paper. CCS-HOG processes the color images in the cross-channel domain to measure the correlation. Specifically, the proposed CCS-HOG consists of the following two steps: 1) to find intrinsic correlation between color channels using the cross-channel similarity measure; 2) to represent the features by extracting histograms of oriented gradients in the cross-channel domain. Furthermore, the descriptors which takes both the correlation and local spatial structure into consideration are presented as features. Finally, comprehensive experiments for two different applications including face recognition and kinship verification are conducted to evaluate the performance of the proposed CCS-HOG.

II. Cross-Channel Similarity based HOG (CCS-HOG)

In this section, we first develop the cross-channel similarity measure to obtain the correlation between color channels, and then propose the CCS-HOG model to extract the features from color images. To better understand the features of CCS-HOG, we visualize their magnitudes and angles in cross-channel domain.

A. Cross-Channel Similarity Measure

The color features have been applied into many feature extraction methods. We develop the cross-channel similarity

978-1-7281-4569-3/19/$31.00 ©2019 IEEE

measure to obtain high-order correlations between color channels ($\{R,G,B\}$). For example, assumed two channel matrices X,Y with the same size of $W \times H$ from color channels ($\{R,G,B\}$), the corresponding cross-channel similarity measure (CCSM) can be expressed as:

$$M(X,Y) = \exp(-\frac{|X-Y|^q}{\sigma}) \quad s.t.\ X,Y \in \{R,G,B\}, \tag{1}$$

where q is the parameter for adjusting measure, σ is the parameter for scale normalization. Different parameter q values lead to different measure results between color channels. For example, it will be an exponential measure if we set $q = 1$; the Eq. 1 will become a gaussian measure when we set $q = 2$. The similarity measure largely depends on the parameter q. Furthermore, the cross-channel similarity measure will determine the final CCS-HOG features. In our model, to make model non-redundant, the CCS of *R-G*, *R-B*, and *G-B* are chosen in our model due to the CCS of channel pair *X-Y* is same as pair *Y-X*.

B. CCS-HOG

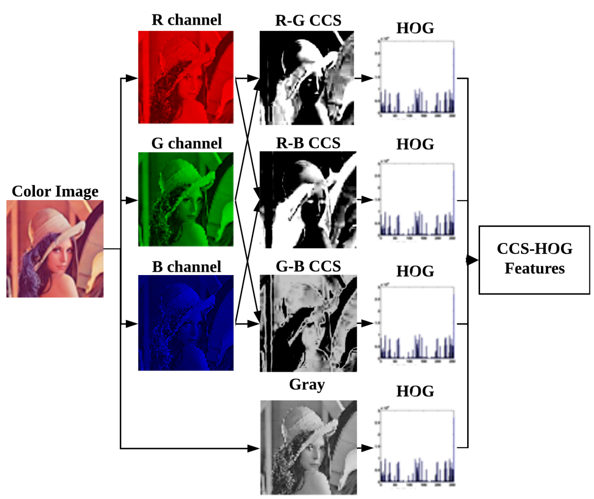

Fig. 1. Flowchart of CCS-HOG model.

The framework of CCS-HOG is illustrated in Fig. 1. Given a color image I with the size of $W \times H$, we acquire the color matrices $\{R,G,B\}$ and the gray matrix denoted as *Gray*. In the first step, the cross-channel similarity measure (CCSM) in Eq. 1 is employed to calculate correlations between $\{R,G,B\}$ color channels to form the three cross-channel similarity matrices. They include *R-G*, *R-B*, and *G-B* similarity matrices . In the second step, we extract the HOG features [8] from these cross-channel similarity matrices and gray matrix. Noted that, we use the l_2-*norm* for block normalization in HOG. Similar to the original HOG in [8], nine bins are set as the parameter of voting strategy. Of course, different numbers of bins will have an influence on performance [8]. Finally, we obtain the CCS-HOG features by concatenating the histograms of oriented gradients extracted from the cross-channel similarity matrices and gray matrix.

C. Visualize CCS-HOG

In traditional HOG [8], the image is filtered by horizontal $[-1\ 0\ 1]$ and vertical $[-1\ 0\ 1]^T$ differential masks. We can obtain the magnitude and phase information from the filtered orthogonal differential signals. Then according to the specific voting strategy in [8], we can obtain the final histograms of oriented gradients from the magnitude and phase information.

To better explain the CCS-HOG, the magnitude and phase feature maps of each pair of cross-channel similarity and gray matrix of a color Lena image are visualized in Fig. 2 and the parameters are set to $\sigma = 255,\ q = 2$.

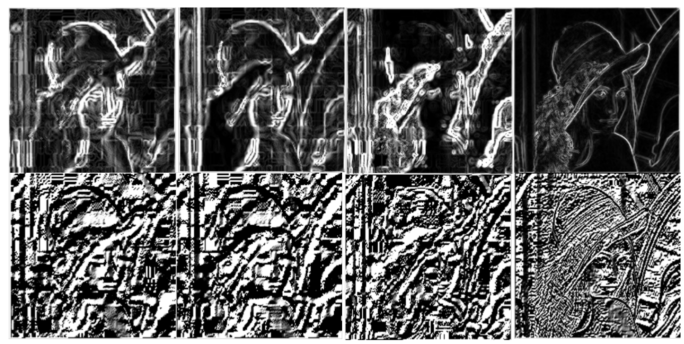

Fig. 2. Visualizing magnitude and phase of of CCS-HOG features, from left to right are R-G, R-B, G-B pairs, and Gray, from top to bottom are corresponding magnitude and phase ($\sigma = 255,\ q = 2$).

III. Experiments

In this section, we conduct experiments of face recognition and kinship verification for color face images to evaluate the performance of the proposed CCS-HOG.

A. Face Recognition

1) Experimental settings: To evaluate the performance of the proposed CCS-HOG, four widely used face databases are

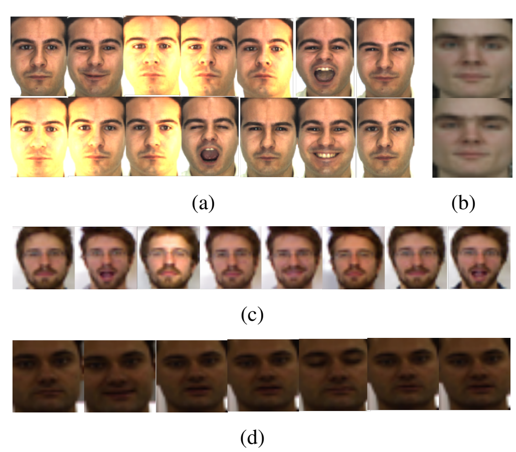

Fig. 3. Aligned and cropped samples of single sample from (a) AR, (b) FERET, (c) EURECOM kinect, (d) CMU PIE.

chosen and several state-of-the-art methods are selected. The face databases include AR face database [17], color FERET database [18], EURECOM kinect database [19], and CMU PIE database [20]. The experiments of face recognition are conducted as [21]. To make experiments more effective, all of the color images in face databases are cropped and resized into 32×32 pixels. The Fig. 3 presents the sample color images from the four databases. We do the experiments of color face recognition by picking out the non-occluded face images from original color face databases. The aforementioned four databases of color face images are organized as follows.

- On AR face database , we employ the non-occluded images in the first session for training and corresponding non-occluded images in another session for testing.
- As for color FERET, the 265 persons in whole database are chosen as subset for experiments. In subset of color FERET, only one of the non-occluded sample from each individual is used to train, another one to test.
- On EURECOM kinect, the non-occluded color images in the first session are used for training and the corresponding non-occluded images in the second session are used for testing.
- As for CMU PIE, the only one sample image is employed to train and the other six color face images are to be composed as test set.

We compare the CCS-HOG with the state-of-the-arts include local binary pattern (LBP) [22], histograms of oriented gradients (HOG) [8], locally encoded transform feature histogram (LETRIST) [23], principal component analysis (PCA) [3], sparse PCA (SPCA) [4], and linear discriminant analysis (LDA) [6]. To conduct a fair comparison, we extend the methods for color images by processing three channels individually and then concatenating them together. Finally, the nearest-neighbor classifier with l_1-*norm* distance is employed for classification.

2) Recognition Performance: The recognition results on the non-occluded color face images are shown in Table I. The best performance of experiments are highlight in bold. On both of four color face databases, CCS-HOG outperforms the other methods although the performance of LETRIST is approaching CCS-HOG on PIE database. Specially, the recognition rate of CCS-HOG is over 10% higher than original HOG in average on most of the face databases.

TABLE I
RECOGNITION PERFORMANCE (%) ON NONOCCLUDED COLOR FACE IMAGES.

Methods	AR	FERET	EU	PIE
LBP	63.14	65.66	47.11	61.03
HOG	76	69.43	62.02	62.25
PCA	77.14	72.45	72.60	54.9
SPCA	75.57	71.7	71.15	50.74
LDA	45.86	73.59	50	54.9
LETRIST	73.71	81.51	61.06	70.58
CCS-HOG	**86.29**	**82.64**	**75**	**71.08**

B. Kinship Verification

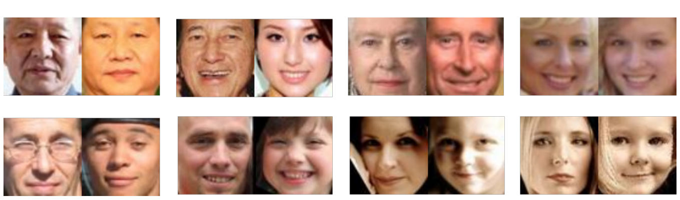

Fig. 4. Aligned and cropped sample kinship pairs, from left to right are the F-S, F-D, M-S and M-D pairs, from top to bottom are the images of KinFaceW-I and KinFaceW-II, respectively.

1) Experimental Settings: The kinship verification is a more challenge problem in case of the aging, gender, and lighting variations between different kinship face images. We employ two widely used kinship face databases, KinFaceW-I (533 pairs) [24] and KinFaceW-II (1000 pairs) [24]. The color kinship face images are acquired under uncontrolled environments. For the KinFaceW-I database, the color face images for a pair are collected from different photos. However, the source of pair images for the KinFaceW-II database are same for the most of pairs. Both of them contain four types of kin relations: Father-Son (F-S), Father-Daughter (F-D), Mother-Son (M-S), and Mother-Daughter (M-D). The KinFaceW-I database contains 156, 134, 116, and 127 pairs of color face images for the four kin relations. The KinFaceW-II database provides 250 pairs of color face images for each kin relation. The aligened and cropped sample color face images of kinship pairs from KinFaceW-I and KinFaceW-II databases are presented in Fig. 5.

In our experiments of kinship verification, the neighborhood repulsed metric learning (NRML) [24] is carried out. For comparison, we choose the state-of-the-art feature extraction methods, namely, LBP [22], HOG [8], weber local descriptor (WLD) [25], local color vector binary pattern (LCVBP) [26], and LETRIST [23]. Similar to experiments of face recognition, the algorithms are derived for color images by concatenating the features of three channels together to conduct a fair comparison.

2) Experimental Results: To evaluate the performance, we choose verification accuracy and ROC curves for comparison. The verification rates of different methods on KinfaceW databases are shown in Table II and Table III. In average, CCS-HOG outperforms other methods on both KinfaceW-I and KinfaceW-II. Specially, the CCS-HOG outperforms the original HOG in all four kin relations from two kinship face databases. The verification rates of CCS-HOG are higher than HOG by around 7% in M-D kin relation pair and 4% in F-S and F-D pairs, 3% in M-S pair in the KinfaceW-I

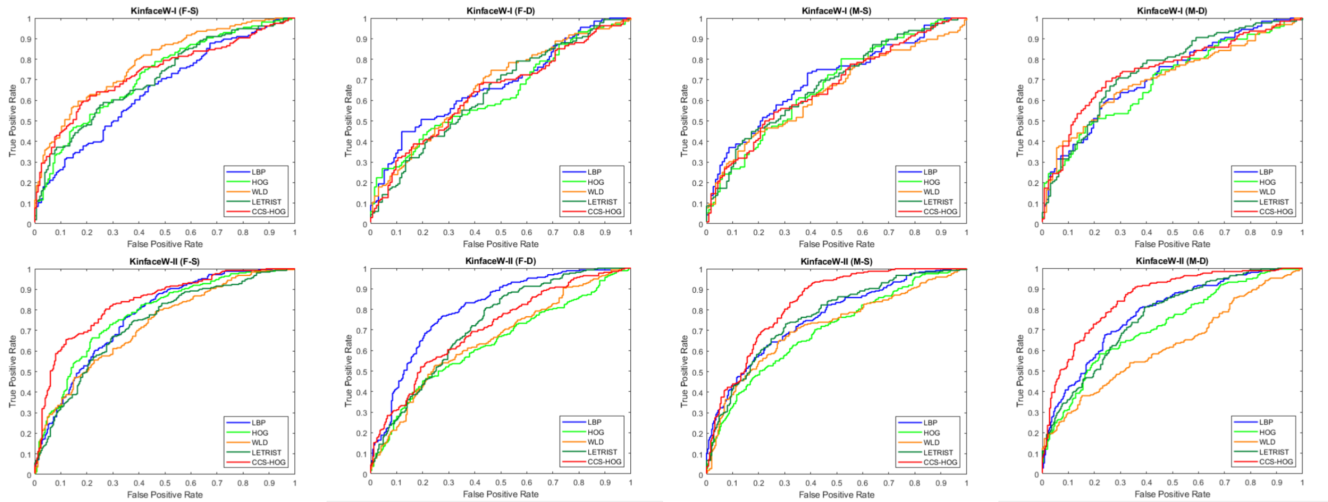

Fig. 5. ROC curves of different methods on KinfaceW databases.

database. As for KinfaceW-II database, the verification rates of CCS-HOG have a significant improvement over HOG with 12% in M-S pair, 10% in M-D pair, 5% in F-S pair, and 4% in F-D pair.

The Fig. 5 plots the ROC curves of different algorithms over five folds on KinfaceW-I and KinfaceW-II databases. In general, the CCS-HOG can obtain the better performance except for the kin relation F-D pair as shown in ROC curves. Therefore, it means F-D pair is more difficult to obtain the correlation similarity. On KinfaceW-I database, CCS-HOG has the better results of ROC curves than HOG on both of four kinship pairs and best performance on M-D kinship pair. As to KinfaceW-II, CCS-HOG can achieve a more promising result for the term of ROC curves. The CCS-HOG attains the best performance of ROC curves in kinship pairs of F-S, M-S, and M-D.

TABLE II
VERIFICATION RATES (%) ON KINFACEW-I.

Methods	KinFace-W-I				
	F-S	F-D	M-S	M-D	Mean
LBP	64.8	**69**	**70.2**	69.7	68.4
HOG	69.6	64.2	67.2	68.1	67.3
WLD	**73.7**	67.2	66.7	70.6	69.6
LCVBP	69.6	66.9	68.5	73	69.5
LETRIST	69.9	65	68.5	73.8	69.3
CCS-HOG	73.4	67.9	**70.2**	**74.81**	**71.6**

TABLE III
VERIFICATION RESULTS (%) ON KINFACEW-II.

Methods	KinFace-W-II				
	F-S	F-D	M-S	M-D	Mean
LBP	73	68	71.8	73.6	71.6
HOG	74.6	64.2	67.6	69.4	69.0
WLD	70	65.6	76.2	63.8	68.9
LCVBP	69.6	**71.6**	75.6	73.4	72.6
LETRIST	71.4	70	72.8	71.6	71.5
CCS-HOG	**79.6**	68.4	**79**	**79.6**	**76.7**

IV. CONCLUSIONS

In this paper, we have developed a cross-channel similarity based histograms of oriented gradients (CCS-HOG) model for color face analysis. In order to extract the effective features from color face images, the cross-channel similarity measure and histograms of gradients are employed in CCS-HOG. Experimental results for face recognition and kinship verification on color images have shown that the proposed method can be widely used and yield better performance than other feature extraction methods.

ACKNOWLEDGMENT

This work was supported in part by the Macau Science and Technology Development Fund under Grant FDCT/189/2017/A3, and by the Research Committee at University of Macau under Grants MYRG2016-00123-FST and MYRG2018-00136-FST.

REFERENCES

[1] Andreas Lanitis, Christopher J. Taylor, and Timothy F. Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 743–756, 1997.

[2] Rajeev Ranjan, Vishal M Patel, and Rama Chellappa, "Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 1, pp. 121–135, 2019.

[3] Svante Wold, Kim Esbensen, and Paul Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987.

[4] Hui Zou, Trevor Hastie, and Robert Tibshirani, "Sparse principal component analysis," *Journal of Computational and Graphical Statistics*, vol. 15, no. 2, pp. 265–286, 2006.

[5] Jian Yang, David D Zhang, Alejandro F Frangi, and Jing-yu Yang, "Two-dimensional pca: a new approach to appearance-based face representation and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004.

[6] Peter N Belhumeur, João P Hespanha, and David J Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, , no. 7, pp. 711–720, 1997.

[7] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, , no. 12, pp. 2037–2041, 2006.

[8] Navneet Dalal and Bill Triggs, "Histograms of oriented gradients for human detection," in *International Conference on Computer Vision & Pattern Recognition*. IEEE Computer Society, 2005, vol. 1, pp. 886–893.

[9] Wenchao Zhang, Shiguang Shan, Wen Gao, Xilin Chen, and Hongming Zhang, "Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition," in *IEEE International Conference on Computer Vision*. IEEE, 2005, vol. 1, pp. 786–791.

[10] Zhiyong Huang, Yuanlong Yu, Jason Gu, and Huaping Liu, "An efficient method for traffic sign recognition based on extreme learning machine," *IEEE Transactions on Cybernetics*, vol. 47, no. 4, pp. 920–933, 2017.

[11] Ling Mao, Mei Xie, Yi Huang, and Yuefei Zhang, "Preceding vehicle detection using histograms of oriented gradients," in *2010 International Conference on Communications, Circuits and Systems*. IEEE, 2010, pp. 354–358.

[12] Minkyu Cheon, Wonju Lee, Changyong Yoon, and Mignon Park, "Vision-based vehicle detection system with consideration of the detecting location," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp. 1243–1252, 2012.

[13] Tejas Indulal Dhamecha, Praneet Sharma, Richa Singh, and Mayank Vatsa, "On effectiveness of histogram of oriented gradient features for visible to near infrared face matching," in *2014 22nd International Conference on Pattern Recognition*. IEEE, 2014, pp. 1788–1793.

[14] Oscar Déniz, Gloria Bueno, Jesús Salido, and Fernando De la Torre, "Face recognition using histograms of oriented gradients," *Pattern Recognition Letters*, vol. 32, no. 12, pp. 1598–1603, 2011.

[15] Omar Oreifej and Zicheng Liu, "Hon4d: Histogram of oriented 4d normals for activity recognition from depth sequences," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 716–723.

[16] Yunsheng Jiang and Jinwen Ma, "Combination features and models for human detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 240–248.

[17] Aleix M Martínez and Avinash C Kak, "Pca versus lda," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 228–233, 2001.

[18] P Jonathon Phillips, Hyeonjoon Moon, Patrick Rauss, and Syed A Rizvi, "The feret evaluation methodology for face-recognition algorithms," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 1997, pp. 137–143.

[19] Rui Min, Neslihan Kose, and Jean-Luc Dugelay, "Kinectfacedb: A kinect database for face recognition," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 11, pp. 1534–1548, 2014.

[20] Terence Sim, Simon Baker, and Maan Bsat, "The cmu pose, illumination, and expression (pie) database," in *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*. IEEE, 2002, pp. 53–58.

[21] Xiaolin Xiao and Yicong Zhou, "Two-dimensional quaternion pca and sparse pca," *IEEE Transactions on Neural Networks and Learning Systems*, 2018.

[22] Timo Ojala, Matti Pietikainen, and Topi Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[23] Tiecheng Song, Hongliang Li, Fanman Meng, Qingbo Wu, and Jianfei Cai, "Letrist: Locally encoded transform feature histogram for rotation-invariant texture classification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 7, pp. 1565–1579, 2018.

[24] Jiwen Lu, Xiuzhuang Zhou, Yap-Pen Tan, Yuanyuan Shang, and Jie Zhou, "Neighborhood repulsed metric learning for kinship verification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 331–345, 2014.

[25] Jie Chen, Shiguang Shan, Chu He, Guoying Zhao, Matti Pietikainen, Xilin Chen, and Wen Gao, "Wld: A robust local image descriptor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1705–1720, 2010.

[26] Seung Ho Lee, Jae Young Choi, Yong Man Ro, and Konstantinos N Plataniotis, "Local color vector binary patterns from multichannel face images for face recognition," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2347–2353, 2012.