Subspace-Based Discrete Transform Encoded Local Binary Patterns Representations for Robust Periocular Matching on NIST's Face Recognition Grand Challenge

Felix Juefei-Xu, *Student Member, IEEE*, and Marios Savvides, *Member, IEEE,*

Abstract—In this paper, we employ several subspace representations (principal component analysis, unsupervised discriminant projection, kernel class-dependence feature analysis, and kernel discriminant analysis) on our proposed discrete transform encoded local binary patterns (DT-LBP) to match periocular region on a large data set such as NIST's face recognition grand challenge (FRGC) ver2 database. We strictly follow FRGC Experiment 4 protocol, which involves 1-to-1 matching of 8014 uncontrolled probe periocular images to 16 028 controlled target periocular images (∼128 million pairwise face match comparisons). The performance of the periocular region is compared with that of full face with different illumination preprocessing schemes. The verification results on periocular region show that subspace representation on DT-LBP outperforms LBP significantly and gains a giant leap from traditional subspace representation on raw pixel intensity. Additionally, our proposed approach using only the periocular region is almost as good as full face with only 2.5% reduction in verification rate at 0.1% false accept rate, yet we gain tolerance to expression, occlusion, and capability of matching partial faces in crowds. In addition, we have compared the best standalone DT-LBP descriptor with eight other state-of-the-art descriptors for facial recognition and achieved the best performance. The two general frameworks are our major contribution: 1) a general framework that employs various generative and discriminative subspace modeling techniques for DT-LBP representation and 2) a general framework that encodes discrete transforms with local binary patterns for the creation of robust descriptors.

Index Terms—Periocular, FRGC, discrete transform, local binary patterns (LBP).

I. INTRODUCTION

OVER the past few years, biometric identification using facial features has gained a lot of prominence. Several papers have suggested novel features and classification techniques that can be used in the case of these modalities in order to improve their performance. Most of these approaches work under the implicit assumption that we are able to capture the entire face of the subject with decent quality. However,
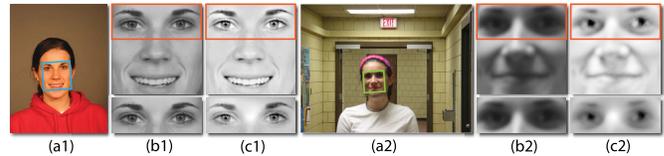
Fig. 1. **Example image from the FRGC generic training set:** (a1, a2) controlled and uncontrolled still, (b1, b2) cropped full face and periocular region with no illumination preprocessing, and (c1, c2) multi-scale retinex (MSR).
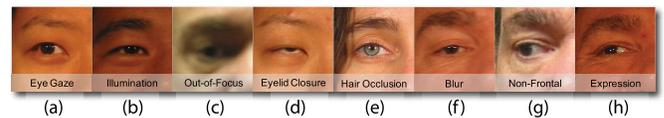

Fig. 2. **Non-ideal eye region variations in the FRGC database [1]:** (a) eye gaze, (b) illumination, (c) out-of-focus, (d) eyelid closure, (e) hair occlusion, (f) blur, (g) non-frontal, and (h) expression.

there are many real-world scenarios where only a partial face is captured or instances when only the eye region of a face is visible, especially for the cases of uncooperative and non-cooperative subjects. Additionally, matching subjects using only the periocular region has great potential in solving "matching faces in the crowd" problem.

In such cases, it is critical to investigate how well facial sub-regions can perform in robust biometric verification. Specifically, we consider the periocular region of the face, which is rich in texture information - eyebrows, eye folds, eyelid contours, etc. which could all vary in shape, size and color. Biologically and genetically speaking, more complex structure means more "coding processing" going on with fetal development, and therefore more proteins and genes involved in the determination of appearance. One can speculate that this is why the periocular region should be the most important facial area for distinguishing people.

Robust periocular based biometric recognition can lead to very useful applications, for example, identifying criminals wearing masks, where only the eye region is exposed, or in videos where there are many occluded faces, but the eye region is un-occluded and can be processed for feature extraction, as shown in Fig. 3. Another scenario where our periocular recognition would be highly desirable is when we want to identify a person in spite of his/her attempts to spoof their iris pattern as in [2]. Using iris patterns alone in these cases can lead to unauthorized access or detection avoidance. Such

a situation can be avoided by using periocular region based recognition in tandem with iris spoof detection techniques.

In this paper, we employ several subspace representations such as principal component analysis (PCA) [3], unsupervised discriminant projection (UDP) [4], kernel class-dependence feature analysis (KCFA) [5], [6], and kernel discriminant analysis (KDA) [7] on our proposed discrete transform encoded local binary patterns (DT-LBP) including WHT-LBP, DCT-LBP, DFT-LBP, DHT-LBP, Legendre-LBP, Chebyshev-LBP, Hermite-LBP and Laguerre-LBP to match the periocular region on a large data set such as NIST's Face Recognition Grand Challenge (FRGC) ver2.0 database. We strictly follow **FRGC EXPERIMENT 4** protocol which involves 1-to-1 matching of 8,014 uncontrolled probe periocular images to 16,028 controlled target periocular images (~128 million pair-wise face match comparisons). The performance of periocular region is compared with that of full face with and without different illumination preprocessing scheme. The verification results show that subspace representation on DT-LBP outperforms LBP significantly on periocular region and gains a giant leap from traditional subspace representation on raw pixel intensity.

To the best of our knowledge, the prior work discussed in Section II lack the following: no one has (1) studied large-scale subspace based periocular matching, (2) compared the performance of periocular region under different illumination conditions with full face on large-scale challenging database, (3) created a general framework such as discrete transform encoded local binary patterns (DT-LBP) to make standard LBP a much better feature descriptor with much higher discriminative power, and (4) modeled the subspace representations using the DT-LBP features to boost the recognition performance over subspace modeling on raw pixel intensities. We focus on the aforementioned 4 major issues that previous works are absent of.

The two general frameworks are our major contribution: (1) a general framework that employs various generative and discriminative subspace modeling techniques for DT-LBP representation, and (2) a general framework that encodes discrete transforms with local binary patterns for the creation of robust descriptors.

Rest of this paper is organized as follows: Section II lists several prior work on periocular region recognition. Section III gives a brief description of the database we use for our experiments as well as the preprocessing and normalization schemes. In Section IV, various feature extraction techniques are described which may be employed to extract relevant discriminating information from a given periocular image. Section V describes various subspace modeling technique employed in this work and the reasoning for choose them. Section VI details the experimental setup for the verification experiments and reports results from the experiments. Section VII compares the proposed DT-LBP descriptor with eight other state-of-the-art descriptors for face recognition. Finally we present some conclusions of our work in Section VIII.

II. RELATED WORK

The earliest study on periocular for biometrics identification can be traced back to Park et al. [8] in 2009. They studied the

feasibility of using periocular images of an individual as a biometric trait. Both global and local features were extracted and fusion of these two was conducted. They showed 77% rank-1 identification rate on a rather small database (958 images from 30 subjects). In the following year, more studies were carried out on periocular region. Juefei-Xu et al. [9] were the first to evaluate the performance of periocular biometrics on a big database as the FRGC ver2.0. They proposed various local feature sets for extracting local features in periocular regions and conducted their experiments following **FRGC EXPERIMENT 4** protocol, the most harsh experiment on the FRGC ver2.0 database. Their results showed that even without any subspace training, periocular region with their proposed feature sets can still outperform NIST's baseline using PCA on full face. More detailed analysis were shown in [1]. Also the discriminative power of the eyebrows was also studied [10]. Merkow et al. [11] explored gender identification using only the periocular region, they showed that at least an 85% classification rate is still obtainable using only the periocular region with a database of 936 low resolution images collected from the internet. Similarly, Lyle et al. [12] also studied gender and ethnicity classification using periocular region features. They used a subset of FRGC and obtained a classification accuracy of 93% and 91% for gender and ethnicity respectively. Miller et al. [13] studied the effect of the quality of the periocular images on recognition performance, the uniqueness of texture between different color channels, and texture information present in different color channels. Woodard et al. [14] utilized periocular region appearance cues for biometric identification both on images captured in visible and NIR spectrum while Park et al. [15] studied periocular biometrics in the visible spectrum. Hollingsworth et al. [16] used near-infrared periocular images to identify useful features for recognition, while others fused periocular with iris images for recognition [17]. Some more recent work using the periocular region include expression tolerance [18] and age invariant face recognition [19], and twin identification using the periocular region [20]. Researcher also tackled the problem of hallucinating the entire face using only the periocular region which showed the effectiveness of using the periocular region when the face is occluded [21].

III. DATABASE AND PREPROCESSING

In this section, we detail the components of NIST's FRGC ver2.0 database as well as the preprocessing and normalization steps before the evaluation experiments.

A. Database

NIST's Face Recognition Grand Challenge (FRGC) ver2.0 database [22] is collected at the University of Notre Dame. Each subject session consists of controlled and uncontrolled still images. The controlled full frontal facial images were taken under two lighting conditions under studio setting with two facial expressions. The uncontrolled images were taken in various locations such as hallways, atria, or outdoors under varying illumination settings also with two expressions, smiling and neutral.

The FRGC ver2.0 database has three components: First, the generic **training** set is typically used in the training phase to extract features. It contains both controlled and uncontrolled images of 222 subjects, and a total of 12,776 images. Second, the **target** set represents the people that we want to find. It has 466 different subjects, and a total of 16,028 images. Last, the **probe** set represents the unknown images that we need to match against the **target** set. It contains the same 466 subjects as in target set, with half as many images for each person as in the target set, bringing the total number of probe images to 8,014. All the **probe** subjects that we are trying to identify are present in the **target** set. The harsh illumination variation mentioned earlier heavily affects the performance of all recognition algorithms and cannot be overlooked. Thus, the need for illumination preprocessing to compensate for the illumination effects emerges, as shown in Section III-B.

FRGC EXPERIMENT 4 is the hardest experiment in the FRGC protocol where face images captured in controlled indoor setting are matched against face images captured in uncontrolled outdoor conditions, where harsh lighting conditions considerably alter the appearance of the face image, as shown in Fig. 1.

Fig. 2 shows non-ideal eye-region variations in the FRGC database. Even though images are taken under the same lighting conditions (controlled or uncontrolled), the intra-class variation is still a challenging problem.

One of the latest trends in face recognition community seems to be working on unconstrained dataset such as the LFW [23], with pose variations, occlusions, expression variations, and illumination variations. Though many algorithms have been proposed that can perform fairly well on such datasets, given the complexity of many of these algorithms, it remains unclear as to what underlying objective each of them aim to achieve in the context of unconstrained face matching. Although success on the LFW framework has been very encouraging, there has been a paradigm shift towards the role of such large unconstrained datasets. It has been suggested that the unconstrained face recognition problems can be decoupled into subtasks where one such factor is tackled at a time [24]. Therefore in this work, we focus on a more constrained face recognition paradigm where many such unconstrained factors have been marginalized out already. The findings of this paper can be easily ported towards unconstrained cases where the proposed feature descriptors can further improve the performance of unconstrained face recognition.

B. Illumination Preprocessing

Illumination is the most significant factor that affects facial appearance besides pose variation. In the FRGC ver2.0 database, the target portion contains controlled images and probe portion contains uncontrolled images. So matching uncontrolled probe images to controlled target images is a challenging task. One common way to tackle this problem is to apply some illumination normalization/preprocessing techniques to compensate for the different lighting conditions.

In order to choose an appropriate illumination normalization scheme among so many studies on illumination invariant

face recognition, we start by looking at 22 prevailing illumination normalization techniques: (1) single scale retinex (SSR) [25], (2) multi scale retinex (MSR) [26], (3) adaptive single scale retinex (ASR) [27], (4) homomorphic filtering (HOMO) [28], (5) single scale self quotient image (SSQ) [29], (6) multi scale self quotient image (MSQ) [29], (7) discrete cosine transform (DCT) [30], (8) retina modeling (RET) [31], (9) wavelet (WA) [32], (10) wavelet denoising (WD) [33], (11) isotropic diffusion (IS) [28], (12) anisotropic diffusion (AS) [34], (13) steering filter (SF) [35], (14) non-local means (NLM) [36], (15) adaptive non-local means (ANL) [36], (16) modified anisotropic diffusion (MAS) [35], (17) gradient-faces (GRF) [37], (18) single scale Weberfaces (WEB) [38], (19) multi scale Weberfaces (MSW) [38], (20) large and small scale features (LSSF) [39], (21) Tan and Triggs (TT) [40], and (22) difference of Gaussian filtering (DOG) [35].

Examples of the 22 illumination normalization schemes are shown in Fig. 4 on an input image with harsh illumination from YaleB [41] and YaleB+ database [42]. There are 3 main reasons to employ YaleB/YaleB+ database, rather than the FRGC ver2.0 database, to select the appropriate normalization scheme: (1) YaleB/YaleB+ database has harsher illumination conditions than the FRGC ver2.0, so testing on this database will push the normalization schemes to their limits; (2) images in YaleB/YaleB+ database are very well aligned because images are collected under a "dome camera" setting, while images in the FRGC are taken by different cameras at different times. This is important because we want the least physical shifts among images when computing the reconstruction error; and (3) we do not want to gain knowledge to improve our system on the unseen testing set. It can be seen from Fig. 4 that some of the illumination normalization outputs are visually better than others. This leads to a very important discussion on how to evaluate and eventually select one scheme out of many. There are 2 main criteria to evaluate illumination normalization techniques, one is the performance improvement in face recognition tasks, and the other one is the normalization quality in terms of how close the normalized face image is to the neutrally/uniformly illuminated face in the database. There is a trade-off between the 2 criteria. Usually any normalization scheme is good at one, but not both. We choose the illumination normalization scheme that reconstructs the image best to neutral lighting one, because in this way, the local binary pattern representations will be least affected by the illumination preprocessing method.

Here, we adopt the peak signal-to-noise ratio (PSNR) as the measurement of illumination normalization quality:

$$\text{PSNR} = 10 \log_{10} \left(\frac{255^2}{\frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2} \right) \quad (1)$$

Within each subject class, the PSNR is computed for each illumination normalized periocular image against the neutrally/uniformly illuminated periocular image in the class. The reported PSNR is the result of averaging over all the frontal images in the YaleB and YaleB+ database shown in Table I. In additional to PSNR, we also report the verification rate (VR) at 0.1% false accept rate (FAR) and equal error

Fig. 3. **Examples of scenarios where periocular recognition is desirable.** (a1) A masked suspect in an ATM armed robbery, (a2) Taliban militant wearing a mask, and (a3) a bank robber wearing a mask, where only their periocular regions are visible for biometric identification. (b1) Parading crowds and (b2) gathering crowds at Washington D.C. for President Obama's Inauguration Ceremony, (b3) crowds in subway station in Tokyo, Japan. There are many occluded faces in the crowd where only their periocular regions are visible for biometric identification. (c1) Muslim women wearing veil, (c2) doctors and nurses wearing clinical mask, and (c3) fire fighter wearing air mask. People from certain religious groups and occupations occlude faces where only their periocular regions are visible for biometric identification.
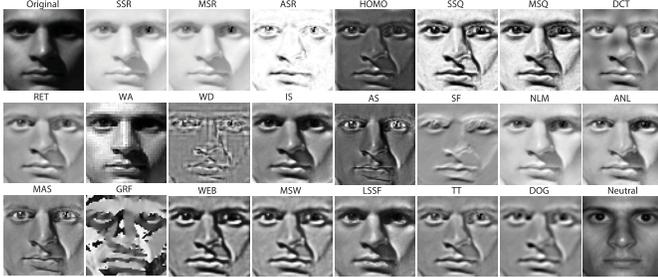

Fig. 4. Comparison of the 22 illumination normalization algorithms on one example face image (top left) with harsh illumination condition from YaleB and YaleB+ Database. Neutrally illuminated face of the same subject is shown in the bottom right. Please zoom in for details.

TABLE I

VR AT 0.001 FAR AND EER FOR YALEB AND YALEB+ EVALUATION.
AVERAGE NORMALIZATION QUALITY IS SHOWN IN PSNR (dB)
(THE HIGHER THE BETTER)

No.	ALGORITHM	VR	EER	PSNR
0	ORIGINAL	0.068	0.484	N/A
1	SSR [25]	0.090	0.351	12.0013
2	MSR [26]	0.079	0.346	**13.8595**
3	ASR [27]	0.210	0.189	5.1167
4	HOMO [28]	0.084	0.432	7.8569
5	SSQ [29]	0.173	0.201	7.7658
6	MSQ [29]	0.158	0.245	8.9509
7	DCT [30]	0.124	0.281	11.9226
8	RET [31]	0.225	0.144	13.1433
9	WA [32]	0.093	0.445	11.3446
10	WD [33]	0.219	0.216	12.6259
11	IS [28]	0.118	0.336	12.5829
12	AS [34]	0.106	0.395	12.9736
13	SF [35]	0.168	0.241	11.4709
14	NLM [36]	0.095	0.331	10.5667
15	ANL [36]	0.095	0.330	11.1804
16	MAS [35]	0.092	0.379	7.9148
17	GRF [37]	0.162	0.303	9.6642
18	WEB [38]	0.156	0.254	11.4052
19	MSW [38]	0.144	0.261	10.9918
20	LSSF [39]	0.100	0.390	12.8664
21	TT [40]	0.226	0.148	13.0195
22	DOG [35]	0.151	0.252	12.8976

rate (EER) of various illumination normalization schemes on YaleB and YaleB+ database.

Based on the results generated from the 22 different illumination normalization schemes we have tried, we find out that multi-scale retinex (MSR) yields the best reconstruction quality with the highest PSNR, as reported in Table I. Hence, we will employ MSR, with the highest PSNR, to compensate for the different lighting conditions. Examples of MSR illumination preprocessing schemes on FRGC ver2.0 database are demonstrated in Fig. 1.

C. Region of Interest Normalization

Rotation and eye coordinate normalization are performed to horizontally align left and right eyes with fixed eye coordinates for all full face frontal images from the target set, probe set

as well as the generic training set so that eyes are registered. Then images are processed for illumination invariance. After the preprocessing, the full face is cropped in size of 128×128 and the periocular region containing both eyes is cropped in size of 50×128. Fig. 1 shows periocular and full face crops.

IV. FEATURE EXTRACTION USING SIGNAL TRANSFORMS

This section gives an overview of individual local feature extraction techniques used in our experiments. First the local binary patterns (LBP) [43] method is introduced with discussion and motivation on using LBP to encode discrete transform coefficients, and then the proposed 8 discrete transform encoded LBP (DT-LBP) such as (1) Walsh-Hadamard transform [44] encoded LBP (WHT-LBP), (2) discrete cosine transform [45] encoded LBP (DCT-LBP), (3) discrete Fourier transform [46] encoded LBP (DFT-LBP), (4) discrete Hartley tranform [47] encoded LBP (DHT-LBP). (5) discrete Legendre polynomial [48]–[50] transform encoded LBP (Legendre-LBP), (6) discrete Chebyshev polynomial [50]–[52] transform encoded LBP (Chebyshev-LBP), (7) discrete Hermite polynomial [50]–[52] transform encoded LBP (Hermite-LBP), and (8) discrete Laguerre polynomial [50]–[52] transform encoded LBP (Laguerre-LBP) are also discussed. Implementation details can be found in corresponding subsections. Also, the last four discrete polynomial transform encoded LBP are described in the Appendix for the sake of compactness.

A. Local Binary Patterns

We start by formulating the traditional LBP operator first introduced by Ojala et al. [43]. Ever since then it has been widely used in biometrics for face recognition [53], face detection [54], facial expression recognition [55], gender classification [56], and iris recognition [57]. Recently, Pietikäinen et al. [58] have summarized the state-of-the-art LBP in the field of computer vision.

The basic idea of this approach is demonstrated in Fig. 5. All neighbors that have values higher than the value of the center pixel are given value 1 and 0 otherwise. The binary numbers associate with the neighbors are then read sequentially to form an binary bit string. The equivalent of this binary number (usually converted to decimal) may be assigned to the center pixel and it may be used to characterize the local texture.

LBP texture for center point (x_c, y_c) can be represented as: $LBP(x_c, y_c) = \sum_{n=0}^{L-1} s(i_n - i_c)2^n$, where i_n denotes the
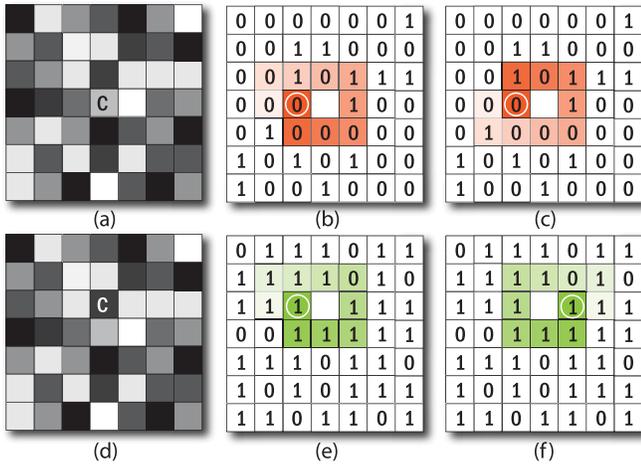
Fig. 5. **Example of the LBP encoding scheme**: (a) 7×7 neighborhood with the thresholding center denoted as C , (b,c) 2 ordering schemes of (a), (d) 7×7 neighborhood with different thresholding center, and (e, f) 2 possible ordering schemes of (d). Different choices of the thresholding center and ordering schemes result in different LBP code.

intensity of the nth surrounding pixel, i_c denotes the intensity of the center pixel, L is the length of the sequence, and $s = 1$ if $i_n \geq i_c$, otherwise, $s = 0$. In the case of a $N \times N$ neighborhood, there are $N^2 - 1$ surrounding pixels, so the bit string is of length $N^2 - 1$.

1) Using LBP to Encode Discrete Transform: In addition to using LBP on pixel intensity for feature extraction, we propose the use of LBP to encode discrete transform coefficients, which is both intuitively reasonable and applicable. The earliest idea of using LBP to encode various discrete transforms can be found in [9].

For example, LBP can be post-applied to encode the DCT coefficients. When DCT is conducted on an image patch, DCT coefficients are obtained. One can directly use all the coefficients as a feature vector, thus representing this image patch in the frequency domain or the transform domain. But from our experience and preliminary results, using these discrete transform coefficients as feature does not yield satisfactory face recognition performance and can easily be beaten by the simplest LBP feature. This is partially due to the fact that the discrete transform coefficients may have a large dynamic range, e.g., in DCT, coefficients associated with lower frequencies tend to have larger values than those associated with higher frequencies. If the dynamic range is not well regulated, the formulation of such features is far from good. Also, directly using discrete transform coefficients as feature is redundant.

Is there some way to remove the redundancy representation of directly using the coefficients while maintaining high discriminative power? The answer is yes. We have discovered that the relative quantities among discrete transform coefficients are more important than their absolute values. So we decide to use LBP to capture the intrinsic spatial characteristics of these coefficients and form an encoding scheme. After the image patch ($3 \times 3, 5 \times 5, 7 \times 7$) is transformed, we apply LBP to encode the discrete transform coefficients using only one decimal number. By doing so, we not only remove the

redundancy (using only 1 number to represent the whole $N \times N$ image patch), but also maintain much higher discriminative power.

What's more, the DT-LBP are expected to greatly outperform standard LBP because what DT-LBP encode is already the second-level feature (coefficients from signal transforms) in frequency domain while LBP encodes the first-level feature (raw pixel intensities) in spatial domain. The frequency-domain representation has much richer and more discriminative information than spatial-domain representation, enabling DT-LBP to be a very robust descriptor. Juefei-Xu et al. have shown that discrete transform encoded local binary patterns (DT-LBP) significantly outperform LBP with no subspace modeling [9], [59]. In Section VI-A, we will continue to explore the performance of DT-LBP and LBP using generative subspace representation and discriminative subspaces.

In addition to normal discrete transforms, we propose to include discrete polynomial transforms such as discrete Legendre polynomial transform, discrete Chebyshev polynomial transform, discrete Hermite polynomial transform, and discrete Laguerre polynomial transform. We first calculate polynomials from low to high orders and then compute its moments in image patches. These polynomials are served as basis and the moments are the projections onto these basis. This is novel way to view the discrete polynomials and they will be included in our discrete transform family for experiments later on.

2) Why to Not Use Uniform LBP: The discrete transform coefficient space is not as smooth as the image intensity space. Uniform LBP constrains that the bit transition happens at most twice in the binary pattern. Ojala et al. claim that most of the LBP in natural images are uniform, but this is not true in discrete transform domain. Moreover, frequent transition is considered as an asset when it comes to encoding because given a pivot/center point, more transitions occurred in the binary code means the pivot is chosen such that coefficients greater than and less than the pivot value are more balanced.

3) Varying Radius: We can extend the methodology to circular or square neighborhoods of increasing radii as demonstrated in Fig. 5. This leads to a multi-resolution representation of the local texture pattern. As the radius of the neighborhood increases, the number of neighbors increases and this may lead to large value when the bit string is converted into a decimal number. To avoid getting such a large decimal value, one can choose fixed number of active neighbors among all the neighbors according to some selection scheme, e.g., choose neighbors only at certain angles and at certain radii [58], or one can vary the base forming the decimal number.

4) Varying Base: We propose to vary the base used for forming the decimal number from the bit string. Instead of using base 2 for conversion as is universally adopted [58], we can use fractional bases (e.g., 1.5, 0.87) or other integer bases (e.g., 3, 4). Unleashing the restriction of using only base 2 for decimal conversion, we can achieve much more diversity when it comes to LBP encoding.

5) Varying Pivot/Center: In the case of 3×3 neighborhood, the center pixel for thresholding neighbors is usually the

physical center of the neighborhood. We propose to vary the center in a larger neighborhood as shown in Fig. 5. Each pivot (thresholding center) gives different bit string, so varying the center will also provide much more diversity.

6) Varying Ordering: If the neighborhood size and the thresholding center are both fixed, different ordering of the neighbors (or the weighting of each bit) gives different decimal outputs. We propose to vary the ordering of the neighbors as shown in Fig. 5 and eventually select an optimal ordering.

In order to achieve the optimal local binary pattern with specific ordering, we propose to borrow the idea from genetic algorithm (GA) [60]. Instead of crossover and mutation in GA, we perform permutation for each new generation. For example, when the neighborhood is 3×3, there are 8 neighbors, and thus the length of bit string is 8. The corresponding weight $w = [2^0 \ 2^1 \ 2^2 \ 2^3 \ 2^4 \ 2^5 \ 2^6 \ 2^7] = [1 \ 2 \ 4 \ 8 \ 16 \ 32 \ 64 \ 128]$ when base is 2. We permutate the weight in each generation and the fitness function is simply the cross validation performance on a subset of the CMU Multi-PIE database [61]. There is limitation of course. When the bit string is short, brute force search is still manageable (total weight permutations for 8-bit string from 3×3 neighborhood is $(3^2 - 1)! = 8! = 40,320$), but when the bit string is long, brute force search becomes impossible (total weight permutations for 48-bit string from 7×7 neighborhood is $(7^2 - 1)! = 48! \approx 1.24 \times 10^{61}$), so we adopt random selection under this circumstance. After certain generations (we have iterated for 10,000 generations), an optimal ordering can be achieved that gives the best performance.

7) Use Histogram Features?: Most of the implementations in [58] utilize the histogram of the LBP image as the feature. We choose not to use the histogram, because histogram quantization does not work when neighborhood sizes increase. Imagine the neighborhood is of size 7×7, and this will lead to a decimal number as large as 2^{49}. In this case, histogram binning is no longer applicable. Additionally, histograms do not exploit the spatial layout of the proposed features. This is why we instead use the LBP/DT-LBP feature as shown in Fig. 8 for subspace modeling to be discussed in Section VI-A.

8) Bit String or Decimal Number?: If the bit string is kept as the feature, the ordering or weights would not be needed. The conversion to decimal number dramatically reduce the feature dimensionality, e.g., when neighborhood size is 7×7, each thresholding center pixel (may not be the physical center) is assigned 48-bit feature and after decimal conversion, the LBP featured image is of the same size as original image, shown in Fig. 8.

9) Circular or Square Neighbors?: The earliest LBP [43] uses circular neighborhood, and the pixel values are bilinearly interpolated whenever the sampling point is not in the center of a pixel. Many implementations in [58] also follow this scheme. Bilinear interpolation make sense when used upon pixel values where intensity of pixel changes continuously. But in our experiments, LBP will be applied to encode the coefficients from discrete transforms, where bilinear interpolation on transform coefficients will
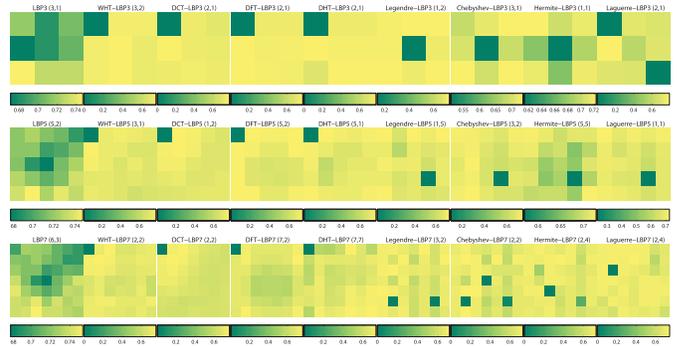

Fig. 6. Selection of pivot point (shown in bracket) for each LBP/DT-LBP feature with 3 different neighborhood sizes. Light colors correspond to high VR.

bring unnecessary noise. So we choose to adopt square neighborhood.

10) Elite Local Binary Pattern: So far, we have seen many variabilities in formulating the DT-LBP feature such as selection of pivot point, ordering, base, and so forth. We have determined that among these variables, selection of pivot point is the most crucial one and greatly affects the recognition performance, while ordering and base variation does not play an important role (from 10,000 evolutions of the genetic algorithm). This is somehow expected because ordering and base variation are by-product from converting the bit-string to decimal number, while selection of pivot is the fundamental in both bit-string and decimal representations.

We carried out preliminary experiments on the CMU Multi-PIE database [61] for evaluating how the selection of pivot will affect the recognition performance. Images are chosen from session 1 with 249 subjects. We use 2 frontal images per subject (1 neutral expression and 1 smiling) under uniform lighting conditions, a total of 498 images, to build our small database. We intentionally choose to use a database other than the FRGC database for model selection and parameter tuning because we do not want to gain knowledge to improve our system on the testing set.

We shown in Fig. 6 the verification rate at 0.1% false accept rate on the small database from the CMU Multi-PIE choosing each of the location as the pivot point on 3×3, 5×5, and 7×7 neighborhood sizes. For each LBP/DT-LBP to be used for the remainder of this paper will follow the pivot selection as shown in Fig. 6.

From Fig. 6, we can see that some coefficients of the DT-LBP always yields low VR when selected as the pivot point. This is because in discrete transforms, some frequency components usually have large values, and when selected as pivot point, the bit string would be less informative (e.g., all 0s). It's also very interesting to see that for LBP, the pivot point that yields the highest VR is not the geometric center of the neighborhood.

B. Walsh-Hadamard Transform Encoded LBP

In this subsection, we introduce the Walsh-Hadamard transform encoded local binary patterns (WHT-LBP) which uses the LBP to encode Walsh-Hadamard transform [44] coefficients.
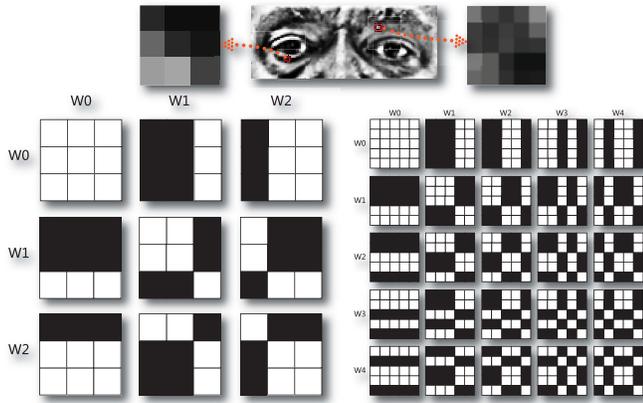
Fig. 7. Example of Walsh-Hadamard transform: expand each 3×3 neighborhood (left) in the periocular image using 9 Walsh basis images or 5×5 neighborhood (right) using 25 Walsh basis images. The coefficients are treated as features assigned to the center pixel.
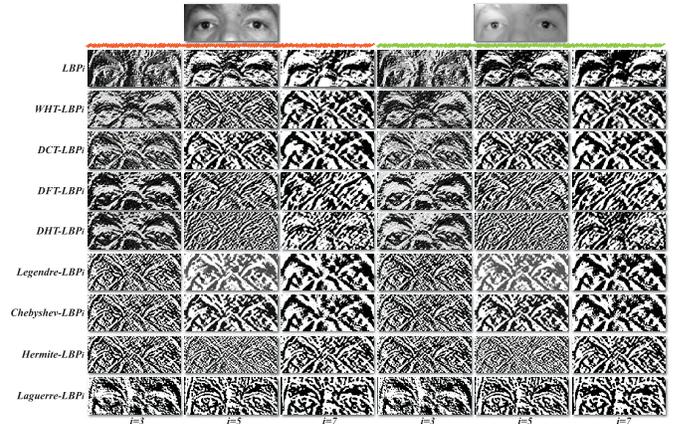

Fig. 8. Examples of the periocular images with the LBP/DT-LBP features using 3 different kernel sizes. The left 3 columns are under no illumination preprocessing, and the right 3 columns are under the MSR illumination preprocessing.

Local image characteristics can be captured by convolution filters such as Walsh masks [62]. The Walsh functions may be used to construct a complete and orthonormal basis in terms of which any matrix of a certain size may be expanded. In our experiments, we use Walsh masks to approximate Walsh-Hadamard transform.

The Walsh matrices consist of two binary patterns, 1s and -1s. Depending on the kernel size, those two patterns are used to represent the local structure of an image. If "local" means a 5×5 kernel window, the Walsh filters we use must correspond to the 5 sample long discrete versions of the Walsh function. Each set of these filters expands the 5×5 image patch in terms of a complete basis of elementary images. The Walsh function:

$$W_{2j+q}(t) = (-1)^{\lfloor \frac{j}{2} \rfloor + q}[W_j(2t) + (-1)^{j+q}W_j(2t-1)], \quad (2)$$

where $\lfloor \frac{j}{2} \rfloor$ means the integer part of $j/2$, q is either 0 or 1. We then sample each function at the integer points only, so we produce the following vectors for size 5:

$$\begin{cases} W_0^T = (1, 1, 1, 1, 1) \\ W_1^T = (-1, -1, -1, 1, 1) \\ W_2^T = (-1, -1, 1, 1, -1) \\ W_3^T = (1, 1, -1, 1, -1) \\ W_4^T = (1, -1, 1, 1, -1) \end{cases} \quad (3)$$

All possible combinations of Walsh vectors can be used to produce 2D basis images. As shown in Fig. 7, there are 9 possible patterns using 3×3 neighborhood and 25 possible patterns using 5×5 neighborhood. We can expand an image in terms of Walsh elementary functions of even size:

$$g = WfW^T, \quad (4)$$

where the transformation matrix W, the input image neighborhood f and the output transform g are all of the same size, $N \times N$ (N is even). An even size transformation matrix constructed from Walsh functions is orthogonal thus its inverse is its transpose.

As shown in previous computation, odd-sized Walsh vectors yield an odd-sized Walsh transformation matrix. Such a matrix

is no longer orthogonal. In order to invert Equation 4, we make use of the inverse of W. Then we shall have [62]: $W^{-1}g(W^T)^{-1} = f$. So, we may use the inverse of matrix W to process the image. Then we have: $g = (W^{-1})^T f W^{-1}$.

After Walsh-Hadamard transform, we use LBP to encode the coefficients to form WHT-LBP, and the feature dimensionality can be reduced to be the same as the size of input image as shown in Fig. 8.

C. Discrete Cosine Transform Encoded LBP

In this subsection, we introduce the discrete cosine transform encoded local binary patterns (DCT-LBP) which uses the LBP to encode discrete cosine transform coefficients.

A discrete cosine transform (DCT) [45] is applied on each $N \times N$ neighborhood in the original image and the N^2 coefficients are treated as the features assigning to the center pixel. The source neighborhood is transformed to a linear combination of these N^2 frequency squares which are the combination of horizontal and vertical frequencies as follows:

$$X_{k_1,k_2} = \sum_{m_1=0}^{M_1-1} \sum_{m_2=0}^{M_2-1} x_{m_1,m_2}$$
$$\times \cos\left[\frac{\pi}{M_1}(m_1 + \frac{1}{2})k_1\right]\cos\left[\frac{\pi}{M_2}(m_2 + \frac{1}{2})k_2\right]. \quad (5)$$

Once we have the DCT coefficients, we apply LBP to encode them to formulate DCT-LBP.

D. Discrete Fourier Transform Encoded LBP

In this subsection, we introduce the discrete Fourier transform encoded local binary patterns (DFT-LBP) which is uses the LBP to encode discrete Fourier transform coefficients.

A 2D discrete Fourier transform (DFT) [46] [44] of an input image $f(x, y)$ of size $M \times N$ is commonly defined as:

$$F(u, v) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)e^{-j2\pi(ux/M+vy/N)},$$
$$u = 0, 1, \ldots, M-1, \quad v = 0, 1, \ldots, N-1. \quad (6)$$

Image processing using 2D DFT would result in complex coefficients, both magnitude and phase of the coefficients are important for image reconstruction. Since magnitude determines the contribution of each component while phase determines which components are present [46], we have reason to believe that phase is more important when it comes to image reconstruction using inverse DFT. But, our experiments found out that, the phase components do not have the spatial characteristics as we expected, on the contrary, the magnitude components satisfy the desired intra-coefficient spatial characteristics. So we use LBP to encode the magnitude components of the DFT coefficients to formulate DFT-LBP.

E. Discrete Hartley Transform Encoded LBP

In this subsection, we introduce the discrete Hartley transform encoded local binary patterns (DHT-LBP) which uses the LBP to encode discrete Hartley transform coefficients.

Discrete Hartley transform (DHT) [47] is a real valued transform which produces real output for a real input and possesses same formula for its own inverse. We employ the rounded Hartley transform (RHT) [63], a transform with zero multiplicative complexity. The RHT therefore has computational advantages over the DHT. The spectrum derived from RHT is a good estimation of the (true) Hartley spectrum and therefore it is used to estimate for DHT [63].

The definition and the philosophy behind the RHT can be explained in such way: let $\mathbf{v}$ be an $\mathbf{n}$-dimensional vector with real elements. The DHT establishes a pair of signal vectors $\mathbf{v} \stackrel{\mathcal{H}}{\longleftrightarrow} \mathbf{V}$, where the elements of $\mathbf{V}$ are defined by:

$$V_k \triangleq \sum_{i=0}^{n-1} v_i \mathrm{cas}\left(\frac{2\pi i k}{n}\right), \quad k = 0, 1, \ldots, n-1, \qquad (7)$$

where $\mathrm{cas}(\theta) \triangleq \cos(\theta) + \sin(\theta)$, an abbreviation adopted from Hartley. This transform leads to the definition of Hartley matrix $\mathbf{H}$, whose elements are of the form $h_{i,k} = \mathrm{cas}(\frac{2\pi i k}{n})$. The rounded Hartley matrix is obatined by rounding off elements of Hartley matrix. Thus the rounded Hartley matrix elements $\widetilde{h}_{i,k}$ are defined by

$$\widetilde{h}_{i,k} \triangleq \left[\underbrace{\mathrm{cas}\left(\frac{2\pi i k}{n}\right)}_{h_{i,k}} \right], \quad i, k = 0, 1, \ldots, N-1, \qquad (8)$$

where $[\cdot]$ denotes the roundoff function. For the sake of notation, rounded Hartley matrix is represented by $\widetilde{\mathbf{H}}_n$.

The definition and effects of 2D RHT on standard classical images are also investigated [63]. Original 2D DHT of an $n \times n$ image is defined by:

$$b_{u,v} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} a_{i,j} \mathrm{cas}\left(\frac{ui + vj}{n}\right), \qquad (9)$$

where $a_{i,j}$ are the elements of an image $\mathbf{A}$ and $b_{u,v}$ are the elements of the Hartley transform of $\mathbf{A}$.

Consider $\mathbf{A}$ as the $n \times n$ image matrix. Let us proceed by calculating a temporary matrix $\mathbf{T}$ as follows: $\mathbf{T} = \widetilde{\mathbf{H}}_n \cdot \mathbf{A} \cdot \widetilde{\mathbf{H}}_n$,

where $\widetilde{\mathbf{H}}_n$ is the rounded Hartley matrix of order n. This is equivalent to taking 1D DHT of the rows, and then transform the columns [64].

Establishing that the elements of $\mathbf{T}$ are represented on the form $t_{i,j}$ $(i, j = 0, 1, \ldots, n-1)$, we can consider three new matrices built from the temporary matrix $\mathbf{T}$: $\mathbf{T}^{(c)}$, $\mathbf{T}^{(r)}$ and $\mathbf{T}^{(cr)}$ whose elements are $t_{i,n-j\,(\mathrm{mod}\,n)}$, $t_{n-i\,(\mathrm{mod}\,n),j}$, $t_{n-i\,(\mathrm{mod}\,n),n-j\,(\mathrm{mod}\,n)}$, respectively. Then the RHT of and $n \times n$ image $\mathbf{A}$ is given by:

$$RHT \triangleq \mathbf{T} + \mathbf{T}^{(c)} + \mathbf{T}^{(r)} - \mathbf{T}^{(cr)}. \qquad (10)$$

After the DHT coefficients are acquired, we apply LBP to encode those coefficients to formulate DHT-LBP.

V. DISCRIMINATIVE SUBSPACE MODELING

A. With No Subspace Modeling

For **FRGC EXPERIMENT 4**, we perform our feature extraction techniques, on **probe** set (8,014 uncontrolled images from 466 subjects), and on **target** set (16,028 controlled images from 466 subjects). This experimental setup is pure 1-to-1 comparison in the strictest sense without using the generic data set at all. In total there are ~128 million comparisons.

B. With Subspace Modeling

In our experiments, we use principal component analysis (PCA) [3], unsupervised discriminant projection (UDP) [4], kernel class-dependence feature analysis (KCFA) [5], [6], and kernel discriminant analysis (KDA) [7] as subspace modeling techniques. Among these models, PCA and UDP are unsupervised, and based on our extensive experiments on the FRGC database, unsupervised methods are usually outperformed by supervised methods. That is why we only kernelize supervised methods (KCFA, KDA) and nonlinear subspace methods also outperform linear subspace methods. Moreover, linear subspace models such as linear discriminant analysis (LDA), independent component analysis (ICA) [74], and locality preserving projection (LPP) [66] are also evaluated in our preliminary experiments. But they do not work well with our proposed LBP/DT-LBP features. Some of the reasons are discussed below.

As shown in Fig. 9, the histograms of the LBP/DT-LBP features are heavily skewed towards both ends and highly non-Gaussian. When the neighborhood size is 5×5 and 7×7, the dynamic range of the feature descriptor is fairly large. In order to make LDA work, where data clusters are modeled using only the mean and class-dependent scatter (covariance) matrices, the assumptions of Gaussianity and homoscedasticity of the data has to be satisfied. Apparently, such assumptions are hardly found in the LBP/DT-LBP features. ICA attempts to decompose a multivariate signal into independent non-Gaussian signals. In other word, the level of non-Gaussianity for the separated signals is actually an optimization criterion for ICA. However, since the LBP/DT-LBP features are already highly non-Gaussian, the signal separation capability from ICA is impeded compared to raw pixel features. Also, the statistical independence assumptions of the data is hard to
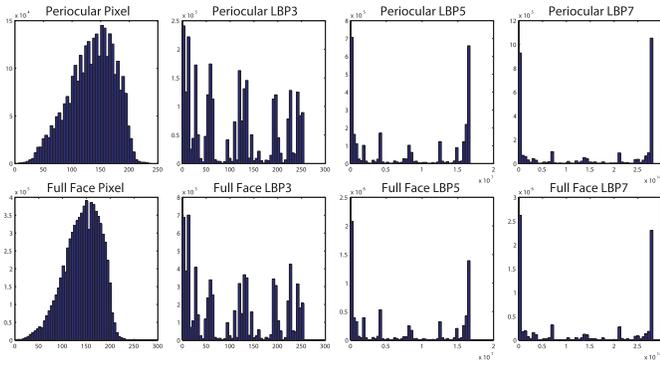
Fig. 9. Histograms of LBP/DT-LBP features versus raw pixel on periocular and full face region.

be satisfied. For LPP, during the adjacency graph forming stages, the distances between pair of data points are computed via the Euclidean norm, and the pair is considered to be neighbors if the distance is less than some pre-determined threshold ϵ. However, the Euclidean norm does not normalize for each dimension of the signal. Thus, it will be dominated by dimensions with very large numbers, leading to a much sparser adjacency graph for a fixed ϵ for the LBP/DT-LBP features, which will hinder the graph embedding procedure in LPP. For the aforementioned reasons, we do not report results on LDA, ICA, and LPP for the sake of brevity.

VI. LARGE-SCALE EXPERIMENTS ON NIST'S FRGC

A. Experimental Setup

In our experiments, we strictly follow NIST's **FRGC EXPERIMENT 4** protocol which involves 1-to-1 matching of 8,014 uncontrolled prove images to 16,028 controlled target images ($\sim$128 million pair-wise face match comparisons). All the raw images go through different illumination pre-processing as mentioned in Section III-B and are aligned and cropped into full face and periocular images as mentioned in Section III-C. We apply each feature extraction technique discussed in Section IV on periocular images as well as on full face images in probe, target, and generic training set accordingly.

Firstly, we carry out our experiments without any subspace training. Secondly, we apply principal component analysis (PCA) method. Thirdly, unsupervised discriminant projection (UDP) method is utilized to model the subspace. Fourthly, we invoke kernel class-dependence feature analysis (KCFA) method. Lastly, kernel discriminant analysis (KDA) is employed as a subspace representation.

Totally, we utilize 5 subspace representation techniques (with no training, PCA, UDP, KCFA, and KDA) to model 10 features (pixel, LBP, WHT-LBP, DCT-LBP, DFT-LBP, DHT-LBP, Legendre-LBP, Chebyshev-LBP, Hermite-LBP and Laguerre-LBP) of 3 different neighborhood sizes on both the periocular and full face images under 2 different illumination preprocessing schemes (uses no illumination preprocessing and MSR).

For PCA, UDP, KCFA, and KDA method, we propose to build subspace on featured (LBP and DT-LBP) images rather

than on raw pixel intensity. In testing phase, we project featured (LBP and DT-LBP) testing images onto those subspaces. We will explore whether LBP still perform better than raw pixel, and whether DT-LBP perform better than LBP after subspace modeling.

For all the experiments we carried out, the normalized cosine distance (NCD) measurement is adopted to compute similarity matrix between target and probe images.

The result of each algorithm is a similarity matrix with the size of $8{,}014 \times 16{,}028$ whose entry SimM_{ij} is the NCD between the feature vector of probe image i and target image j. The performance of each feature extraction technique in each experimental setup under each illumination preprocessing method is analyzed using verification rate (VR) at 0.1% false accept rate (FAR) and the performance of best DT-LBP against LBP on periocular region is compared using receiver operating characteristic (ROC) curves.

B. Experimental Results

In this section, experimental results are reported and followed by quantitative analysis and discussions.

Raw images go through different illumination preprocessing schemes and then are aligned and cropped into full face images with size of 128×128 and periocular images with size of 50×128 as shown in Fig. 1. Proposed feature extraction techniques such as LBP and DT-LBP are applied on both periocular region and full face.

We propose several subspace representations (PCA, UDP, KCFA, and KDA) on LBP/DT-LBP featured images and compare their verification performance with traditional subspace representation on raw pixel intensity. Performance of periocular region from each algorithm is compared with that of corresponding full face.

Table II and Table III show the verification rate (VR) (or equivalently true accept rate) at 0.1% false accept rate (FAR) for **FRGC EXPERIMENT 4**. Experimental results with no illumination preprocessing are listed in Table II and experimental results under MSR illumination preprocessing scheme are shown in Table III. The left 5 columns of the tables show experiments on periocular region and the right 5 columns shows that of full face.

The experiments employ 5 subspace representations (no training (NT), principal component analysis (PCA), unsupervised discriminant projection (UDP), kernel class-dependence feature analysis (KCFA), and kernel discriminant analysis (KDA)), to model 10 feature types (raw pixel, LBP, WHT-LBP, DCT-LBP, DFT-LBP, DHT-LBP, Legendre-LBP, Chebyshev-LBP, Hermite-LBP and Laguerre-LBP). For LBP/DT-LBP, 3 different neighborhood sizes (3×3, 5×5, and 7×7) are evaluated and reported individually with $\star$ showing the best among 3 neighborhood sizes. Scores labeled with $\star\star$ are the best performing LBP and DT-LBP for periocular region and will be used to plot the ROC curves in Fig. 10.

From Table II and Table III, we can see that:
1) Neighborhood size does affect the verification performance. For LBP and each type of the proposed DT-LBP, the best performing neighborhood size may not

TABLE II

VR AT 0.1% FAR FOR FRGC EXPERIMENT 4 ON PERIOCULAR REGION AND FULL FACE WITH NO ILLUMINATION PREPROCESSING

	PERIOCULAR REGION					FULL FACE				
	NT	PCA	UDP	KCFA	KDA	NT	PCA	UDP	KCFA	KDA
Pixel	0.008	0.001	0.002	0.430	0.427	0.009	0.002	0.002	0.471	0.474
LBP₃	0.021	0.058	0.161	0.546	0.478**	0.016	0.040	0.175	0.576	0.500
LBP₅	0.047	0.085	0.178**	0.542	0.424	0.090*	0.117*	0.218*	0.615	0.584
LBP₇	0.052**	0.090**	0.115	0.552**	0.444	0.069	0.107	0.153	0.620*	0.600*
WHT-LBP₃	0.050	0.083	0.153*	0.489	0.183	0.082	0.116	0.199*	0.549	0.262
WHT-LBP₅	0.137*	0.162	0.126	0.475	0.325	0.143	0.149	0.107	0.451	0.297
WHT-LBP₇	0.123	0.166*	0.132	0.566**	0.486*	0.145*	0.176*	0.174	0.596*	0.478*
DCT-LBP₃	0.062	0.128	0.148	0.491	0.152	0.058	0.143	0.181	0.491	0.180
DCT-LBP₅	0.141*	0.181**	0.186**	0.556	0.469	0.154*	0.175*	0.183*	0.566	0.428
DCT-LBP₇	0.100	0.144	0.111	0.558*	0.519*	0.127	0.171	0.173	0.601*	0.525*
DFT-LBP₃	0.047	0.089	0.147*	0.491	0.149	0.050	0.117	0.203*	0.556	0.254
DFT-LBP₅	0.131*	0.150	0.124	0.468	0.313	0.143	0.134	0.096	0.450	0.296
DFT-LBP₇	0.123	0.164*	0.115	0.541*	0.454*	0.152*	0.170*	0.139	0.558*	0.466*
DHT-LBP₃	0.055	0.097	0.162*	0.509	0.169	0.055	0.123*	0.213*	0.562	0.265
DHT-LBP₅	0.096*	0.096	0.087	0.309	0.228	0.086*	0.082	0.057	0.260	0.195
DHT-LBP₇	0.063	0.109*	0.128	0.546*	0.335*	0.071	0.106	0.124	0.585*	0.467*
Legendre-LBP₃	0.044	0.075	0.116	0.455	0.204	0.118	0.124	0.214*	0.533	0.392
Legendre-LBP₅	0.151**	0.162	0.117	0.463	0.355	0.166	0.165	0.114	0.461	0.337
Legendre-LBP₇	0.136	0.164*	0.133*	0.558*	0.548**	0.184*	0.197*	0.189	0.607*	0.486*
Chebyshev-LBP₃	0.056	0.103	0.118	0.447	0.163	0.099	0.155	0.179	0.482	0.265
Chebyshev-LBP₅	0.149*	0.176*	0.181*	0.551	0.502	0.187*	0.195*	0.194*	0.575	0.370
Chebyshev-LBP₇	0.110	0.143	0.106	0.554*	0.547*	0.166	0.192	0.187	0.615*	0.495*
Hermite-LBP₃	0.083	0.125	0.128*	0.474	0.169	0.106	0.143	0.167*	0.483	0.197
Hermite-LBP₅	0.116	0.126	0.098	0.379	0.214	0.123	0.108	0.072	0.351	0.218
Hermite-LBP₇	0.145*	0.168*	0.104	0.496*	0.369*	0.160*	0.159*	0.123	0.488*	0.382*
Laguerre-LBP₃	0.046	0.073	0.124*	0.472	0.217	0.114	0.138	0.211*	0.555	0.366
Laguerre-LBP₅	0.096*	0.095	0.078	0.305	0.231	0.081	0.084	0.056	0.267	0.199
Laguerre-LBP₇	0.066	0.109*	0.111	0.513*	0.345*	0.129*	0.151*	0.133	0.583*	0.382*
Best Features Combined	0.236	0.267	0.282	0.728	0.751	0.249	0.271	0.284	0.746	0.773

TABLE III

VR AT 0.1% FAR FOR FRGC EXPERIMENT 4 ON PERIOCULAR REGION AND FULL FACE WITH MULTI-SCALE RETINEX (MSR) ILLUMINATION PREPROCESSING

	PERIOCULAR REGION					FULL FACE				
	NT	PCA	UDP	KCFA	KDA	NT	PCA	UDP	KCFA	KDA
Pixel	0.013	0.001	0.002	0.425	0.375	0.031	0.017	0.098	0.592	0.558
LBP₃	0.016	0.044	0.119**	0.488	0.417	0.046	0.108	0.197	0.553	0.361
LBP₅	0.038	0.071**	0.115	0.489	0.385**	0.120*	0.132	0.225*	0.602	0.545
LBP₇	0.043**	0.063	0.077	0.496**	0.381	0.116	0.135*	0.167	0.626*	0.580*
WHT-LBP₃	0.041	0.070	0.108	0.442	0.181	0.132	0.131	0.213*	0.523	0.368
WHT-LBP₅	0.150*	0.163	0.119	0.465	0.357	0.163	0.163	0.110	0.445	0.318
WHT-LBP₇	0.134	0.165*	0.130*	0.556**	0.538**	0.190*	0.204*	0.187	0.600*	0.467*
DCT-LBP₃	0.058	0.099	0.114	0.437	0.146	0.111	0.164	0.169	0.468	0.234
DCT-LBP₅	0.148*	0.175*	0.177*	0.550*	0.495	0.194*	0.196	0.188*	0.565	0.358
DCT-LBP₇	0.108	0.143	0.107	0.546	0.534*	0.176	0.200*	0.188	0.610*	0.487*
DFT-LBP₃	0.035	0.064	0.098	0.437	0.136	0.118	0.141	0.207*	0.541	0.340
DFT-LBP₅	0.132*	0.148	0.117*	0.464	0.235	0.151	0.145	0.095	0.444	0.227
DFT-LBP₇	0.127	0.161*	0.113	0.530*	0.390*	0.175*	0.189*	0.145	0.559*	0.394*
DHT-LBP₃	0.043	0.070	0.113*	0.460	0.164	0.126	0.147	0.210*	0.547	0.346
DHT-LBP₅	0.095*	0.094	0.077	0.303	0.224	0.082	0.082	0.054	0.263	0.188
DHT-LBP₇	0.067	0.104*	0.107	0.503*	0.337*	0.140*	0.157*	0.131	0.572*	0.363*
Legendre-LBP₃	0.102	0.127	0.131*	0.466	0.165	0.156	0.137	0.166*	0.465	0.187
Legendre-LBP₅	0.122	0.134	0.098	0.392	0.258	0.132	0.121	0.085	0.360	0.251
Legendre-LBP₇	0.167**	0.186**	0.125	0.544*	0.454*	0.174*	0.182*	0.153	0.546*	0.414*
Chebyshev-LBP₃	0.058	0.103	0.120	0.447	0.163	0.098	0.155	0.179	0.482	0.267
Chebyshev-LBP₅	0.150*	0.175*	0.181**	0.550	0.500	0.187*	0.195*	0.194*	0.578	0.368
Chebyshev-LBP₇	0.111	0.144	0.106	0.552*	0.547*	0.167	0.193	0.185	0.615*	0.496*
Hermite-LBP₃	0.083	0.126	0.126*	0.474	0.168	0.109	0.143	0.168*	0.482	0.194
Hermite-LBP₅	0.118	0.125	0.099	0.375	0.214	0.121	0.107	0.072	0.348	0.216
Hermite-LBP₇	0.145*	0.165*	0.107	0.496*	0.369*	0.159*	0.157*	0.123	0.489*	0.383*
Laguerre-LBP₃	0.044	0.076	0.125*	0.473	0.215	0.111	0.138	0.211*	0.553	0.367
Laguerre-LBP₅	0.096*	0.096	0.076	0.303	0.232	0.081	0.085	0.054	0.267	0.198
Laguerre-LBP₇	0.068	0.111*	0.114	0.514*	0.347*	0.129*	0.151*	0.131	0.583*	0.381*
Best Features Combined	0.242	0.261	0.279	0.733	0.747	0.252	0.279	0.289	0.769	0.777

be consistent across various subspace representations, but usually larger neighborhood size gives better performance.

2) The best verification rate of LBP and DT-LBP (WHT-LBP, DCT-LBP, DFT-LBP, DHT-LBP, Legendre-LBP, Chebyshev-LBP, Hermite-LBP and Leguerre-LBP) is significantly better than that of pixel, across all the subspace representations except for KDA which assumes the input data to be Gaussian. Since LBP/DT-LBP images are no longer Gaussian, KDA on raw pixel sometimes performs better than on LBP/DT-LBP.

3) With less than 40% the size of a full face ($\frac{50 \times 128}{128 \times 128} = 0.3906$), periocular region still perform outstandingly. Under certain circumstances, periocular region performs even better than full face. Table IV shows the average drop in VR at 0.1% FAR on periocular region from full face using each subspace representation. We find that periocular region, on average, has only 2.5% drop in VR at 0.1% FAR compared to full face, yet we gain tolerance to expression, occlusion and capability of matching partial faces in crowds.

4) The best DT-LBP is always better than the best LBP on periocular region with and without illumination pre-processing.

5) The best DT-LBP are WHT-LBP, Legendre-LBP, DCT-LBP and Chebyshev-LBP across various subspace representations.

6) When best DT-LBP features are combined, a remarkable 75.1% VR at 0.1% FAR is achievable using only the periocular region on the FRGC database.

7) The total genuine matches in the FRGC ver2.0 database is 407,352, so 0.01 (1%) improvement in VR
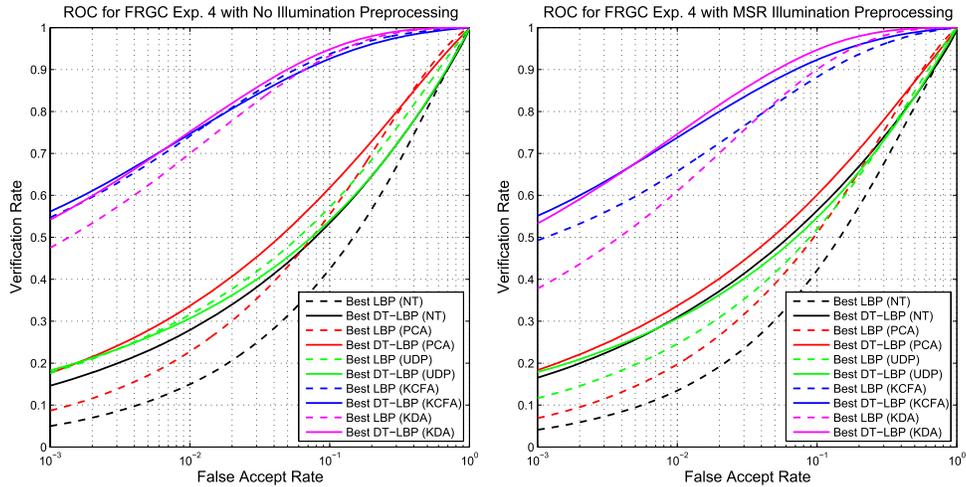
Fig. 10.　ROC curves for FRGC Experiment 4 on periocular region only with no illumination preprocessing (left) and MSR (right). Comparisons of the best LBP and the best DT-LBP for each subspace representation are shown.

at 0.1% FAR means 4,073 more entries are being correctly verified, which is significant. Table V shows the improvement of the best DT-LBP over the best LBP on periocular region for all subspace representations. From the results in Table V, we can achieve as high as 15.3% improvement in VR at 0.1% FAR using DT-LBP feature, a giant leap over standard LBP feature.

8) Table V also shows the cross-illumination preprocessing comparisons in VR on the best LBP and the best DT-LBP. We can see that when shifting from no illumination preprocessing to MSR, the best LBP always drops in performance and the best DT-LBP only drops at certain subspace representations. It is important to notice that DT-LBP has 66.86% lower standard deviation in VR across illumination preprocessing schemes than LBP. This leads to a significant finding that DT-LBP feature on average is much less affected by different illumination normalization schemes, thus more robust than LBP feature.

Fig. 10 shows the receiver operating characteristics (ROC) curves of the best LBP and the best DT-LBP for each subspace representation under the two illumination preprocessing schemes.

VII. COMPARISON TO THE STATE-OF-THE-ART DESCRIPTORS FOR FACE RECOGNITION

In this section, we compare the proposed DT-LBP descriptors to several state-of-the-art descriptors for face recognition. They are: (1) Gabor [75], (2) SIFT [76], (3) HOG [77], (4) GLOH [78], (5) SURF [79], (6) BRIEF [80], (7) DAISY [81], and (8) BRISK [82]. We have seen that a majority of the latest efforts on unconstrained face recognition (usually on the Labeled Faces in the Wild (LFW) database) greatly reply on locating accurate and dense facial landmarks. High dimensional descriptors are then extracted from each of the landmarks to achieve good performances such as in [83]–[87]. The need for accurate facial landmarks is due to the fact that off-angle faces are being matched to each other under unconstrained scenarios. In this work, since we are focusing

TABLE IV

AVERAGE VR DROP AT 0.1% FAR FOR FRGC EXPERIMENT 4 ON PERIOCULAR REGION FROM FULL FACE

	NT	PCA	UDP	KCFA	KDA	Average
No Illum.	0.0225	0.0149	0.0267	0.0275	0.0336	**0.0250**
MSR	0.0423	0.0311	0.0430	0.0393	0.0332	0.0378

TABLE V

VR AT 0.1% FAR FOR FRGC EXPERIMENT 4 ON PERIOCULAR REGION FROM THE BEST LBP AND THE BEST DT-LBP FEATURES

	NT	PCA	UDP	KCFA	KDA	
	NO ILLUMINATION PREPROCESSING					
Best LBP	0.052	0.090	0.178	0.552	0.478	
Best DT-LBP	0.151	0.181	0.186	0.566	0.548	
Improvement	0.099	0.091	0.008	0.014	0.070	
	MSR ILLUMINATION PREPROCESSING					
Best LBP	0.043	0.071	0.119	0.496	0.385	
Best DT-LBP	0.167	0.186	0.181	0.556	0.538	
Improvement	0.124	0.115	0.062	0.060	**0.153**	
	CROSS ILLUM. PREPROC. COMPARISON					Std. σ
Best LBP	-0.009	-0.019	-0.059	-0.056	-0.093	0.0338
Best DT-LBP	0.016	0.005	-0.005	-0.010	-0.010	0.0112

on facial descriptors, we will constrain ourselves on frontal or near-frontal facial images. Thus, we can ignore facial landmarks because the facial image are very well aligned already.

In the following experiments, we will again utilize the **FRGC EXPERIMENT 4** for the evaluation of various state-of-the-art descriptors. Without any training stage, we want to match the probe set (8,014 images) directly to the target set (16,028 images). For each periocular image, we extract the best stand-alone DT-LBP feature: Legendre-LBP5, as well as the eight other facial descriptors mentioned above. In this evaluation, no dimensionality reduction is involved, and therefore for the sake of fair comparison, we will keep equivalent dimensions among all descriptors. For example, the DT-LBP descriptors have the same dimension as the input image, and we can keep all 6,400 dimensions (the periocular image is $50 \times 128 = 6400$). However, for SIFT descriptor, at each keypoint, a 128-dim feature vector is extracted. Therefore, we uniformly sample the image for 50 keypoint locations and perform SIFT at each keypoint. The resulting feature vector is the concatenation of the feature vectors at all keypoints $(50 \times 128 = 6400)$. This applies to many other descriptors.

TABLE VI
VR at 0.1% FAR for FRGC Experiment 4 on Periocular
Region From the Legendre-LBP5 and Eight
State-of-the-Art Descriptors

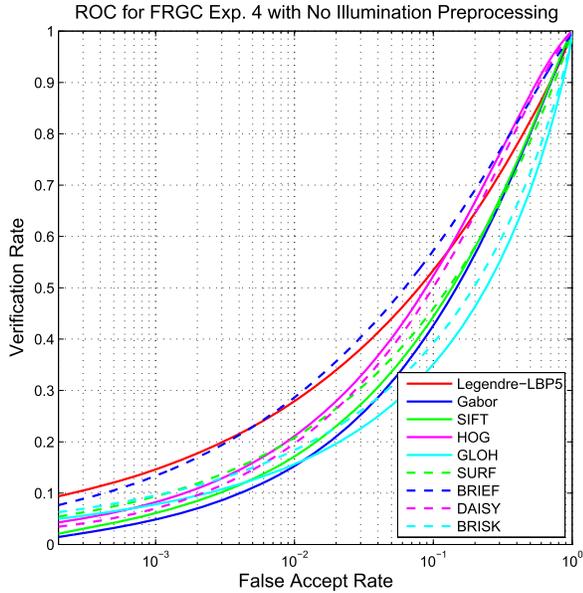| Descriptors	VR at 0.1% FAR
Legendre-LBP5	**0.151**
Gabor [75]	0.048
SIFT [76]	0.063
HOG [77]	0.083
GLOH [78]	0.077
SURF [79]	0.093
BRIEF [80]	0.135
DAISY [81]	0.070
BRISK [82]	0.096

Fig. 11. ROC curves for FRGC Experiment 4 on periocular region only with no illumination preprocessing. Comparisons of Legendre-LBP5 with eight state-of-the-art descriptors are shown.

Table VI shows the VR at 0.1% FAR for this comparative evaluation. The proposed Legendre-LBP5 solidly outperforms other facial descriptors. Fig. 11 shows the ROC curves for this evaluation. Many descriptors that are performing well for the tasks such as scene matching, and object classification may not be best suited for face recognition. This can be attributed to the lack of strong edge gradients and corners on human faces.

VIII. Conclusions

The two general frameworks are our major contribution: (1) a general framework that employs various generative and discriminative subspace modeling techniques for DT-LBP representation, and (2) a general framework that encodes discrete transforms with local binary patterns for the creation of robust descriptors.

To be more specific, we have shown that the subspace representation (PCA, UDP, KCFA, and KDA) on both local binary patterns (LBP) and our proposed discrete transform encoded local binary patterns (DT-LBP) features significantly outperform traditional subspace representations on raw pixel intensities. Additionally, experimental results show that subspace representations on the best DT-LBP feature significantly outperforms the best standard LBP with **15.3%** improvement in VR at 0.1% FAR on periocular region and full face with different illumination preprocessing schemes. We also find out

that with less than 40% the size of a full face, periocular region performs amazingly well compared to full face, with only **2.5%** drop in VR at 0.1% FAR. Yet, we are able to gain tolerance to illumination, expression, occlusion, and capability of matching partial faces in crowds. Under certain conditions, periocular region is even better than full face in terms of verification performance under non-ideal scenarios. In addition, we have compared the best stand-alone DT-LBP descriptor with eight other state-of-the-art descriptors for facial recognition and achieved the best performance, which re-confirmed the robustness and efficacy of the proposed DT-LBP feature for face recognition.

APPENDIX A

DISCRETE POLYNOMIAL TRANSFORM ENCODED LBP

In this section, we visit various polynomial families that we harnessed in coining the discrete polynomial transform encoded LBP. Specifically, they are Legendre polynomial, Chebyshev polynomial, Hermite polynomial, and Laguerre polynomial.

A. Discrete Legendre Polynomial Transform Encoded LBP

In this subsection, we introduce the discrete Legendre polynomial transform encoded local binary patterns (Legendre-LBP) which uses the LBP to encode discrete Legendre polynomial transform coefficients.

Legendre polynomials [48]–[50], also called Legendre functions are solutions of of Legendre differential equations:

$$\frac{d}{dx}\left[(1-x^2)\frac{d}{dx}P_n(x)\right] + n(n+1)P_n(x) = 0 \quad (11)$$

The Legendre differential equation can be solved using standard power series method. The solutions form a polynomial sequence, the Legendre polynomials. Using Rodrigues's formula, each nth-degree Legendre polynomial $P_n(x)$ can be expressed as:

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\left[(x^2 - 1)^n\right] \quad (12)$$

By applying the general Leibniz rule for repeated differentiation, the $P_n(x)$ can also be defined as the coefficients in a Taylor series expansion:

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n \quad (13)$$

In order to obtain the recursive relationship for Legendre polynomials, Equation 13 is differentiated on both sides with respect to t, and we have:

$$\frac{x-t}{\sqrt{1-2xt+t^2}} = (1-2xt+t^2)\sum_{n=1}^{\infty} nP_n(x)t^{n-1} \quad (14)$$

Combining with Equation 13, and equating the coefficients of power of t gives Bonnet's recursion formula:

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x) \quad (15)$$

with first two terms being $P_0(x) = 1$ and $P_1(x) = x$ pursuant to Taylor series expansion.

We can also obtain the explicit representation of Legendre polynomials by induction from Bonnet's recursion formula:

$$P_n(x) = \sum_{k=0}^{n} (-1)^k \binom{n}{k}^2 \left(\frac{1+x}{2}\right)^{n-k} \left(\frac{1-x}{2}\right)^k \quad (16)$$

Legendre polynomials are orthogonal on the $x \in [-1, 1]$ interval:

$$\int_{-1}^{1} P_m(x)P_n(x)dx = \frac{2}{2n+1}\delta_{mn} \quad (17)$$

where δ_{mn} denotes the Kronecker delta.

The 2D Legendre polynomial transform (Legendre moments) [88]–[90] of an image intensity function $f(x, y)$ is defined as:

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4} \int_{-1}^{1}\int_{-1}^{1} P_m(x)P_n(x)f(x, y)dxdy \quad (18)$$

The region of definition of a Legendre polynomial is the interior of the unit circle, so any rectangle image will have to be scaled to the region $-1 < x, y < 1$. For a rectangle image of size $M \times N$, we have the following discrete Legendre polynomial transform:

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4MN} \\ \times \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} P_m(-1+\frac{2x}{M-1})P_n(-1+\frac{2y}{N-1})f(x, y) \quad (19)$$

After discrete Legendre polynomial transform, we apply LBP to encode discrete Legendre polynomial transform coefficients and form Legendre-LBP. Example is demonstrated in Fig. 8.

B. Discrete Chebyshev Polynomial Transform Encoded LBP

In this subsection, we introduce the discrete Chebyshev polynomial transform encoded local binary patterns (Chebyshev-LBP) which uses the LBP to encode discrete Chebyshev polynomial transform coefficients.

Chebyshev polynomials [50]–[52] are solutions to the Chebyshev differential equations:

$$(1 - x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} + n^2y = 0 \quad (20)$$

and

$$(1 - x^2)\frac{d^2y}{dx^2} - 3x\frac{dy}{dx} + n(n+2)y = 0 \quad (21)$$

for the polynomials of the first and second kind respectively. Chebyshev polynomials of the first kind are more important in image processing, so in our implementation we will use the first kind.

We can obtain the explicit representation of Chebyshev polynomials by:

$$T_n(x) = \frac{n}{2} \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \frac{(n-k-1)!}{k!(n-2k)!}(2x)^{n-2k} \quad (22)$$

Similar to the Legendre polynomials, the Chebyshev polynomials can be expressed using recursive relationship:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad (23)$$

with $T_0(x) = 1$ and $T_1(x) = x$.

Chebyshev polynomials are orthogonal with respect to the weight $\frac{1}{\sqrt{1-x^2}}$ on the $x \in (-1, 1)$ interval:

$$\int_{-1}^{1} T_n(x)T_m(x)\frac{dx}{\sqrt{1-x^2}} = \begin{cases} 0 & n \neq m, \\ \pi & n = m = 0, \\ \pi/2 & n = m \neq 0. \end{cases}$$

The discrete Chebyshev polynomials can be written as:

$$T_n(x) = n! \sum_{k=0}^{n} (-1)^{n-k} \binom{N-1-k}{n-k}\binom{n+k}{n}\binom{x}{k} \quad (24)$$

which satisfy the orthogonality relation:

$$\sum_{x=0}^{N-1} T_n(x)T_m(x) = \varrho(n, N)\delta_{mn} \quad (25)$$

with the normalizing factor:

$$\varrho(n, N) = \frac{N(N^2 - 1^2)(N^2 - 2^2)\cdots(N^2 - n^2)}{2n+1} \\ = (2n)!\binom{N+n}{2n+1}, \quad n = 0, 1, \ldots, N-1 \quad (26)$$

The recurrence formula for discrete Chebyshev polynomials is:

$$T_{n+1}(x) = \frac{(2n+1)(2x-N+1)}{n+1}T_n(x) - \frac{n(N^2-n^2)}{n+1}T_{n-1}(x) \quad (27)$$

Then the discrete Chebyshev polynomial transform [90]–[92] for rectangle images of size $M \times N$ is:

$$\tau_{pq} = \frac{1}{\varrho(p, M)\varrho(q, N)} \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} T_m(x)T_n(y)f(x, y) \quad (28)$$

Finally, we utilize LBP to encode discrete Chebyshev polynomial transform coefficients and formulate Chebyshev-LBP. Example is shown in Fig. 8.

C. Discrete Hermite Polynomial Transform Encoded LBP

In this subsection, we introduce the discrete Hermite polynomial transform encoded local binary patterns (Hermite-LBP) which uses the LBP to encode discrete Hermite polynomial transform coefficients.

The probabilists' Hermite polynomials [50]–[52] are solutions to the differential equation:

$$\frac{d}{dx}\left(e^{\frac{-x^2}{2}}\frac{du}{dx}\right) + \lambda e^{\frac{-x^2}{2}}u = 0 \quad (29)$$

where λ is a constant with the boundary conditions that u to be bounded polynomially at infinity. This differential equation can be rewritten as an eigenvalue problem (Hermite equation): $u'' - xu' = -\lambda u$, and both will lead to:

$$H_n(x) = (-1)^n e^{x^2/2}\frac{d^n}{dx^n}e^{-x^2/2} \quad (30)$$

Hermite polynomials can be explicitly represented as:

$$H_n(x) = n! \left(\frac{1}{\sqrt{2}}\right)^n \sum_{\ell=0}^{n/2} \frac{(-1)^{n/2-\ell}}{(2\ell)!(n/2-\ell)!}(\sqrt{2}x)^{2\ell} \quad (31)$$

for odd values of n, and for even values the representation follows:

$$H_n(x) = n! \left(\frac{1}{\sqrt{2}}\right)^n \sum_{\ell=0}^{(n-1)/2} \frac{(-1)^{(n-1)/2-\ell}(\sqrt{2}x)^{2\ell+1}}{(2\ell+1)!((n-1)/2-\ell)!} \quad (32)$$

We can also use the following recurrence relationship to express Hermite Polynomials:

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x) \quad (33)$$

with $H_0(x) = 1$ and $H_1(x) = 2x$.

The nth degree Hermite polynomials are orthogonal with respect to the weight function e^{-x^2}:

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)e^{-x^2/2}dx = \sqrt{2\pi}n!\delta_{nm} \quad (34)$$

Then the discrete Hermite polynomial transform [93]–[95] for rectangle images of size $M \times N$ is:

$$\hbar_{pq} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} H_p(\frac{x}{\sigma})H_q(\frac{y}{\sigma})f(x,y) \quad (35)$$

where σ is the standard deviation of the Gaussian function.

We can obtain the Hermite-LBP by post-applying LBP encoding scheme on discrete Hermite polynomial transform coefficients. Example is shown in Fig. 8.

D. Discrete Laguerre Polynomial Transform Encoded LBP

In this subsection, we introduce the discrete Laguerre polynomial transform encoded local binary patterns (Laguerre-LBP) which uses the LBP to encode discrete Laguerre polynomial transform coefficients.

The generalized Laguerre polynomials [50]–[52] are solutions of generalized Laguerre's equation:

$$x\frac{d^2y}{dx^2} + (\alpha+1-x)\frac{dy}{dx} + ny = 0 \quad (36)$$

and can also be expressed by the Rodrigues' formula:

$$L_n^{(\alpha)}(x) = \frac{x^{-\alpha}e^x}{n!}\frac{d^n}{dx^n}(e^{-x}x^{n+\alpha}) \quad (37)$$

Simple Laguerre polynomials are recovered by setting $\alpha = 0$.

The explicit representation of generalized Laguerre polynomials of degree n using Leibniz's theorem for differentiation of a product[1] to Rogrigues' formula is:

$$L_n^{(\alpha)}(x) = \sum_{i=0}^{n}(-1)^i \binom{n+\alpha}{n-i}\frac{x^i}{i!} \quad (38)$$

Lagurre polynomials can be expressed using the following recurrence relationship:

$$L_n^{(\alpha)}(x) = \left(2+\frac{\alpha-1-x}{n}\right)L_{n-1}^{(\alpha)}(x) - \left(1+\frac{\alpha-1}{n}\right)L_{n-2}^{(\alpha)}(x) \quad (39)$$

with $L_0^{(\alpha)}(x) = 1$ and $L_1^{(\alpha)}(x) = -x + \alpha + 1$.

Laguerre polynomials are orthogonal over $[0, \infty)$ with respect to the measure with the weighting function $x^\alpha e^{-x}$:

$$\int_0^\infty x^\alpha e^{-x} L_n^{(\alpha)}(x)L_m^{(\alpha)}(x)dx = \frac{\Gamma(n+\alpha+1)}{n!}\delta_{nm} \quad (40)$$

Then we can obtain the discrete Laguerre polynomial transform [96], [97] for rectangle images of size $M \times N$ as follows:

$$\ell_{pq} = \frac{1}{\rho(p,M)\rho(q,N)} \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} e^{-x}x^\alpha L_p^\alpha(x)L_q^\alpha(y)f(x,y) \quad (41)$$

where $\rho(p,M) = \sum_{x=0}^{M-1} e^{-x}x^\alpha\{L_p^\alpha(x)\}^2$.

Following the same manner as other discrete polynomial transforms, we can achieve Laguerre-LBP by encoding the discrete Laguerre polynomial transform coefficients with LBP. Example is illustrated in Fig. 8.

---

[1]The generalized Leibniz rule states that if function f and g are n-times differentiable, then the nth derivative of the product fg is given by $(f \cdot g)^{(n)} = \sum_{k=0}^{n}\binom{n}{k}f^{(k)}g^{(n-k)}$.

REFERENCES

[1] F. Juefei-Xu, M. Cha, M. Savvides, S. Bedros, and J. Trojanova, "Robust periocular biometric recognition using multilevel fusion of various local feature extraction techniques," in *Proc. IEEE 17th Int. Conf. Digit. Signal Process., DSP*, 2011.

[2] J. Galbally, J. Fierrez, and J. Ortega-Garcia, "Vulnerabilities in biometric systems: Attacks and recent advances in liveness detection," in *Proc. Spanish Workshop Biometrics*, 2007.

[3] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York, NY, USA: Springer-Verlag, 2002.

[4] J. Yang, D. Zhang, J.-Y. Yang, and B. Niu, "Globally maximizing, locally minimizing: Unsupervised discriminant projection with applications to face and palm biometrics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 650–664, Apr. 2007.

[5] C. Xie, M. Savvides, and B. V. K. V. Kumar, "Redundant class-dependence feature analysis based on correlation filters using FRGC2.0 data," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2005, p. 153.

[6] B. V. K. V. Kumar, M. Savvides, and C. Xie, "Correlation pattern recognition for face recognition," *Proc. IEEE*, vol. 94, no. 11, pp. 1963–1976, Nov. 2006.

[7] C. Xie and B. V. K. V. Kumar, "Comparison of kernel class-dependence feature analysis (KCFA) with kernel discriminant analysis (KDA) for face recognition," in *Proc. 1st IEEE Int. Conf. Biometrics, Theory, Appl., Syst., BTAS*, Sep. 2007, pp. 1–6.

[8] U. Park, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum: A feasibility study," in *Proc. IEEE 3rd Int. Conf. Biometrics, Theory, Appl., Syst., BTAS*, Sep. 2009, pp. 1–6.

[9] F. Juefei-Xu, M. Cha, J. L. Heyman, S. Venugopalan, R. Abiantun, and M. Savvides, "Robust local binary pattern feature sets for periocular biometric identification," in *Proc. 4th IEEE Int. Conf. Biometrics, Theory, Appl., Syst., BTAS*, Sep. 2010, pp. 1–8.

[10] F. Juefei-Xu and M. Savvides, "Can your eyebrows tell me who you are?" in *Proc. 5th Int. Conf. Signal Process. Commun. Syst., ICSPCS*, Dec. 2011, pp. 1–8.

[11] J. Merkow, B. Jou, and M. Savvides, "An exploration of gender identification using only the periocular region," in *Proc. 4th IEEE Int. Conf. Biometrics, Theory, Appl., Syst., BTAS*, Sep. 2010, pp. 1–5.

[12] J. R. Lyle, P. E. Miller, S. J. Pundlik, and D. L. Woodard, "Soft biometric classification using periocular region features," in *Proc. 4th IEEE Int. Conf. Biometrics, Theory, Appl., Syst., BTAS*, Sep. 2010, pp. 1–7.

[13] P. E. Miller, J. R. Lyle, S. J. Pundlik, and D. L. Woodard, "Performance evaluation of local appearance based periocular recognition," in *Proc. 4th IEEE Int. Conf. Biometrics, Theory, Appl., Syst., BTAS*, Sep. 2010, pp. 1–6.

[14] D. L. Woodard, S. J. Pundlik, J. R. Lyle, and P. E. Miller, "Periocular region appearance cues for biometric identification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, CVPRW*, Jun. 2010, pp. 162–169.

[15] U. Park, R. Jillela, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 96–106, Mar. 2011.

[16] K. Hollingsworth, K. W. Bowyer, and P. J. Flynn, "Identifying useful features for recognition in near-infrared periocular images," in *Proc. 4th IEEE Int. Conf. Biometrics, Theory Appl. Syst., BTAS*, Sep. 2010, pp. 1–8.

[17] D. L. Woodard, S. Pundlik, P. Miller, R. Jillela, and A. Ross, "On the fusion of periocular and iris biometrics in non-ideal imagery," in *Proc. 20th Int. Conf. Pattern Recognit., ICPR*, Aug. 2010, pp. 201–204.

[18] F. Juefei-Xu and M. Savvides, "Unconstrained periocular biometric acquisition and recognition using cots PTZ camera for uncooperative and non-cooperative subjects," in *Proc. IEEE Workshop Appl. Comput. Vis.*, Jan. 2012, pp. 201–208.

[19] F. Juefei-Xu, K. Luu, M. Savvides, T. Bui, and C. Suen, "Investigating age invariant face recognition based on periocular biometrics," in *Proc. Int. Joint Conf. Biometrics, IJCB*, Oct. 2011, pp. 1–7.

[20] F. Juefei-Xu and M. Savvides, "An augmented linear discriminant analysis approach for identifying identical twins with the aid of facial asymmetry features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, CVPRW*, Jun. 2013, pp. 56–63.

[21] F. Juefei-Xu, D. K. Pal, and M. Savvides, "Hallucinating the full face from the periocular region via dimensionally weighted K-SVD," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, CVPRW*, Jun. 2014.

[22] P. J. Phillips *et al.*, "Overview of the face recognition grand challenge," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., CVPR*, vol. 1. Jun. 2005, pp. 947–954.

[23] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Dept. Comput., Univ. Massachusetts Amherst, Amherst, MA, USA, Tech. Rep. 07-49, Oct. 2007.

[24] J. Z. Leibo, Q. Liao, and T. Poggio, "Subtasks of unconstrained face recognition," in *Proc. Int. Joint Conf. Comput. Vis. Imag. Comput. Graph. Theory Appl., VISAPP*, 2014.

[25] D. J. Jobson, Z.-U. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. Image Process.*, vol. 6, no. 3, pp. 451–462, Mar. 1997.

[26] D. J. Jobson, Z.-U. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 965–976, Jul. 1997.

[27] Y. K. Park, S. L. Park, and J. K. Kim, "Retinex method based on adaptive smoothing for illumination invariant face recognition," *Signal Process.*, vol. 88, no. 8, pp. 1929–1945, 2008.

[28] G. Heusch, F. Cardinaux, and S. Marcel, "Lighting normalization algorithms for face verification," in *Proc. IDIAP-Com*, Mar. 2005.

[29] H. Wang, S. Z. Li, Y. Wang, and J. Zhang, "Self quotient image for face recognition," in *Proc. Int. Conf. Image Process., ICIP*, vol. 2. Oct. 2004, pp. 1397–1400.

[30] W. Chen, M. J. Er, and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 36, no. 2, pp. 458–466, Apr. 2006.

[31] N.-S. Vu and A. Caplier, "Illumination-robust face recognition using retina modeling," in *Proc. 16th IEEE Int. Conf. Image Process., ICIP*, Nov. 2009, pp. 3289–3292.

[32] S. Du and R. Ward, "Wavelet-based illumination normalization for face recognition," in *Proc. IEEE Int. Conf. Image Process., ICIP*, vol. 2. Sep. 2005. pp. 954–957.

[33] T. Zhang *et al.*, "Multiscale facial structure representation for face recognition under varying illumination," *Pattern Recognit.*, vol. 42, no. 2, pp. 251–258, Feb. 2009.

[34] R. Gross and V. Brajovic, "An image preprocessing algorithm for illumination invariant face recognition," in *Proc. 4th Int. Conf. Audio Video-Based Biometric Pers. Authentication*, vol. 2. Jun. 2003, pp. 1–10.

[35] V. Štruc and N. Pavešić, "Performance evaluation of photometric normalization techniques for illumination invariant face recognition," Internal report, LUKS, 2010.

[36] V. Štruc and N. Pavešić, "Illumination invariant face recognition by non-local smoothing," in *Proc. Int. Conf. BIOID Multicomm.*, Sep. 2009, pp. 1–8.

[37] T. Zhang, Y. Y. Tang, B. Fang, Z. Shang, and X. Liu, "Face recognition under varying illumination using gradientfaces," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2599–2606, Nov. 2009.

[38] B. Wang, W. Li, W. Yang, and Q. Liao, "Illumination normalization based on Weber's law with application to face recognition," *IEEE Signal Process. Lett.*, vol. 18, no. 8, pp. 462–465, Aug. 2011.

[39] X. Xie, W.-S. Zheng, J. Lai, P. C. Yuen, and C. Y. Suen, "Normalization of face illumination based on large-and small-scale features," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1807–1821, Jul. 2011.

[40] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1635–1650, Jun. 2010.

[41] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001.

[42] K.-C. Lee, J. Ho, and D. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 5, pp. 684–698, May 2005.

[43] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, 1996.

[44] D. F. Elliott and K. R. Rao, *Fast Transforms: Algorithms, Analyses, Applications*. San Francisco, CA, USA: Academic, 1982.

[45] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithm, Advantages, Applications*. San Francisco, CA, USA: Academic, 1990.

[46] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989.

[47] R. N. Bracewell, "Discrete hartley transform," *J. Opt. Soc. Amer.*, vol. 73, no. 12, pp. 1832–1835, Dec. 1983.

[48] S. Omachi and M. Omachi, "Fast template matching with polynomials," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2139–2149, Aug. 2007.

[49] G. Li and C. Wen, "Legendre polynomials in signal reconstruction and compression," in *Proc. 5th IEEE Conf. Ind. Electron. Appl., ICIEA*, Jun. 2010, pp. 1636–1640.

[50] R. A. El. Attar, *Special Functions and Orthogonal Polynomials*. Raleigh, NC, USA: Lulu Press, 2006.

[51] J. Baik, T. Kriecherbauer, K. McLaughlin, and P. D. Miller, *Discrete Orthogonal Polynomials*. Princeton, NJ, USA: Princeton Univ. Press, 2007.

[52] G. Sansone, *Orthogonal Functions*. New York, NY, USA: Dover, 2004.

[53] H. Zhang and D. Zhao, "Spatial histogram features for face detection in color images," in *Proc. 5th Pacific Rim Conf. Multimedia Adv. Multimedia Inform. Process., PCM*, vol. 1. 2004, pp. 377–384.

[54] H. Jin, Q. Liu, H. Lu, and X. Tong, "Face detection using improved LBP under Bayesian framework," in *Proc. 3rd Int. Conf. Image Graph.*, Dec. 2004, pp. 306–309.

[55] C. Shan, S. Gong, and P. W. McOwan, "Robust facial expression recognition using local binary patterns," in *Proc. IEEE Int. Conf. Image Process., ICIP*, vol. 2. Sep. 2005, pp. 370–376.

[56] N. Sun, W. Zheng, C. Sun, C. Zou, and L. Zhao, "Gender classification based on boosting local binary pattern," in *Proc. 3rd Int. Symp. Neural Netw., ISNN*, vol. 3972. 2006, pp. 194–201.

[57] Z. Sun, T. Tan, and X. Qiu, "Graph matching iris image blocks with local binary pattern," in *Proc. Int. Conf. Adv. Biometrics, ICB*, 2006, pp. 366–373.

[58] M. Pietikainen, G. Zhao, A. Hadid, and T. Ahonen, *Computer Vision Using Local Binary Patterns*. New York, NY, USA: Springer-Verlag, 2011.

[59] F. Juefei-Xu, M. Cha, M. Savvides, S. Bedros, and J. Trojanova, "Robust periocular biometric recognition using multi-level fusion of various local feature extraction techniques," in *Proc. 17th Int. Conf. Digit. Signal Process., DSP*, Jul. 2011, pp. 1–8.

[60] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.

[61] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-pie," in *Proc. 8th IEEE Int. Conf. Autom. Face Gesture Recognit.*, Sep. 2008, pp. 1–8.

[62] M. Petrou and P. G. Sevilla, *Dealing with Texture*. New York, NY, USA: Wiley, 2006.

[63] R. J. de Sobral Cintra, H. M. de Oliveira, and C. O. Cintra, "Rounded Hartley transform: A quasi-involution," in *Proc. Int. Telecommun. Symp., ITS*, 2002.

[64] R. C. Gonzalez and R. E. Woods, *Digital Imaging Processing*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2007.

[65] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit., CVPR*, Jun. 1991, pp. 586–591.

[66] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.

[67] B. V. K. V. Kumar, A. Mahalanobis, and R. D. Juday, *Correlation Pattern Recognition*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[68] A. Mahalanobis, B. V. K. V. Kumar, and D. Casasent, "Minimum average correlation energy filters," *Appl. Opt.*, vol. 26, no. 17, pp. 3633–3640, Sep. 1987.

[69] J. Figue and P. Réfrégier, "Optimality of trade-off filters," *Appl. Opt.*, vol. 32, no. 11, pp. 1933–1935, Apr. 1993.

[70] M. Savvides, B. V. K. V. Kumar, and P. K. Khosla, "'corefaces'—Robust shift invariant PCA based correlation filter for illumination tolerant face recognition," in *Proc. Conf. Comput. Vis. Pattern Recognit., CVPR*, vol. 2. Jul. 2004, pp. 834–841.

[71] M.-H. Yang, "Kernel eigenfaces vs. kernel fisherfaces: Face recognition using kernel methods," in *Proc. 5th IEEE Int. Conf. Autom. Face Gesture Recognit.*, May 2002, pp. 215–220.

[72] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[73] S. Mika, G. Rätsch, J. Weston, B. Schölkopf, and K. Mullers, "Fisher discriminant analysis with kernels," in *Proc. IEEE Signal Process. Soc. Workshop Neural Netw. Signal Process.*, Aug. 1999, pp. 41–48.

[74] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York, NY, USA: Wiley, 2001.

[75] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 467–476, Apr. 2002.

[76] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[77] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., CVPR*, vol. 1. Jun. 2005, pp. 886–893.

[78] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[79] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Surf: Speeded up robust features," *Comput. Vis. Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[80] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief: Binary robust independent elementary features," in *Proc. Eur. Conf. Comput. Vis., ECCV*, 2010, pp. 778–792.

[81] E. Tola, V. Lepetit, and P. Fua, "DAISY: An efficient dense descriptor applied to wide-baseline stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 815–830, May 2010.

[82] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *Proc. IEEE Int. Conf. Comput. Vis., ICCV*, Nov. 2011, pp. 2548–2555.

[83] K. Simonyan, O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Fisher vector faces in the wild," in *Proc. Brit. Mach. Vis. Conf., BMVC*, 2013.

[84] D. Chen, X. Cao, F. Wen, and J. Sun, "Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification," in *Proc. IEEE Comput. Vis. Pattern Recognit., CVPR*, 2013, pp. 3025–3032.

[85] H. Li, G. Hua, Z. Lin, J. Brandt, and J. Yang, "Probabilistic elastic matching for pose variant face verification," in *Proc. IEEE Comput. Vis. Pattern Recognit., CVPR*, Jun. 2013, pp. 3499–3506.

[86] Q. Cao, Y. Ying, and P. Li, "Similarity metric learning for face recognition," in *Proc. IEEE Int. Conf. Comput. Vis., ICCV*, Dec. 2013, pp. 2408–2415.

[87] O. Barkan, J. Weill, L. Wolf, and H. Aronowitz, "Fast high dimensional vector multiplication face recognition," in *Proc. IEEE Int. Conf. Comput. Vis., ICCV*, Dec. 2013, pp. 1960–1967.

[88] C. H. Teh and R. T. Chin, "On image analysis by the methods of moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 4, pp. 496–513, Jul. 1988.

[89] P.-T. Yap and R. Paramesran, "An efficient method for the computation of Legendre moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 12, pp. 1996–2002, Dec. 2005.

[90] J. Flusser, T. Suk, and B. Zitova, *Moments and Moment Invariants in Pattern Recognition*. New York, NY, USA: Wiley, 2009.

[91] R. Mukundan, S. H. Ong, and P. A. Lee, "Image analysis by tchebichef moments," *IEEE Trans. Image Process.*, vol. 10, no. 9, pp. 1357–1364, Sep. 2001.

[92] H. Shu, H. Zhang, B. Chen, P. Haigron, and L. Luo, "Fast computation of tchebichef moments for binary and grayscale images," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3171–3180, Dec. 2010.

[93] Y. Bo and M. Dai, "Image analysis by Gaussian–Hermite moments," *Signal Process.*, vol. 91, no. 10, pp. 2290–2303, Oct. 2011.

[94] K. M. Hosny, "Fast computation of accurate Gaussian-Hermite moments for image processing applications," *Digit. Signal Process.*, vol. 22, no. 3, pp. 476–485, 2012.

[95] Y. Wu and J. Shen, "Properties of orthogonal Gaussian-Hermite moments and their applications," *EURASIP J. Appl. Signal Process.*, vol. 4, pp. 588–599, Jan. 2005.

[96] H. Qjidaa, "Image reconstruction by Laguerre moments," in *Proc. Int. Symp. Commun., Control Signal Process.*, 2006.

[97] S. S. Mizrahi and D. Galetti, "Laguerre moments and generalized functions," *J. Phys. A, Math. General*, vol. 35, no. 15, pp. 3535–3546, 2002.

Felix Juefei-Xu (S'10) received the B.S. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009, and the M.S. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2010, where he is currently pursuing the Ph.D. degree. He is currently with the CMU CyLab Biometrics Center. His research is mainly focused on biometrics, pattern recognition, machine learning, computer vision, and image processing.

Marios Savvides (S'03–M'04) is the Founder and Director of the Biometrics Center with Carnegie Mellon University (CMU), Pittsburgh, PA, USA, and an Associate Research Professor with the Department of Electrical and Computer Engineering and the Carnegie Mellon CyLab. He is also one of the tapped researchers to form the Office of the Director of National Intelligence, 1st Center of Academic Excellence in Science and Technology. His research is mainly focused on developing algorithms for robust face and iris biometrics, and pattern recognition, machine vision, and computer image understanding for enhancing biometric systems performance. He is on the program committee on several biometric conferences such as the IEEE International Conference on Biometrics: Theory, Applications and Systems, the IEEE International Conference on Pattern Recognition, the SPIE Biometric Identification, the IEEE AutoID, and others, and has also organized and co-chaired the 2008 Robust Biometrics Understanding the Science and Technology Conference. He has authored and co-authored over 170 journal and conference publications, including several book chapters in the area of biometrics, and served as the Area Editor of *Encyclopedia of Biometrics* (Springer). His achievements include leading the research and development in CMU's past participation at NIST's Face Recognition Grand Challenge in 2005 and also in NIST's Iris Challenge Evaluation. He holds over 20 patent applications in the area of biometrics, and was a recipient of the CMU's 2009 Carnegie Institute of Technology Outstanding Faculty Research Award.