Weight-optimal Local Binary Patterns

Felix Juefei-Xu and Marios Savvides

Carnegie Mellon University

Abstract. In this work, we have proposed a learning paradigm for obtaining weight-optimal local binary patterns (WoLBP). We first reformulate the LBP problem into matrix multiplication with all the bitmaps flattened and then resort to the Fisher ratio criterion for obtaining the optimal weight matrix for LBP encoding. The solution is closed form and can be easily solved using one eigen-decomposition. The experimental results on the FRGC ver2.0 database have shown that the WoLBP gains significant performance improvement over traditional LBP, and such WoLBP learning procedure can be directly ported to many other LBP variants to further improve their performances.

Keywords: Local Binary Patterns (LBP), Weight-optimal Local Binary Patterns (WoLBP).

1 Introduction

In the field of computer vision and pattern recognition, local binary patterns (LBP) and its variants have been widely used throughout many applications. The LBP has gained its prominence due to its discriminative power and computational simplicity. The simple yet very efficient operator labels the pixels of an image (patch) by thresholding the neighborhood of each pixel and converts the result as a binary number.

The LBP was invented in 1992, with the idea that two-dimensional textures can be described by two complementary local measures: pattern and contrast [21]. By separating pattern information from contrast, invariance to monotonic gray scale changes can be obtained. The first published work using LBP for face recognition is done by Ahonen et al. in 2004 [1], where they divided the face image into several regions from which the LBP features are extracted and concatenated into an enhanced feature vector to be used as a face descriptor. Following this, LBP and its variants have been widely used in the field of biometrics for face recognition [26], face detection [3], facial expression recognition [23], gender classification [24], and iris recognition [25]. Recently, Pietikäinen et al. [21] have summarized the state-of-the-art LBP in the field of computer vision and pattern recognition. More face recognition and analysis related work using LBP variants can be found in [6, 4, 8, 9, 5, 10, 11, 13, 12, 7].

Recently, many efforts are devoted to learning optimal local binary patterns for various applications. For example, Lei et al. [14] (an extended work of [15]) have learned discriminant face descriptor by first learning the discriminant image filters; second, soft determining the optimal neighborhood sampling strategy;

and third, statistically constructing the dominant patterns. Their method is iterative and relies on 2D-LDA type of formulation, which is quite computationally expensive. Shan [22] uses AdaBoost to select discriminative LBP features for gender classification. Liao et al. [17] again applies AdaBoost to select the most effective uniform multi-scale block LBP for enhanced face recognition. Only this time, the computation is done based on average values of block subregions, instead of individual pixels. Maturana et al. [18] consider the following method. Within any square neighborhood given by r, there are $(2r+1)^2-1$ possible pixel comparisons. They wish to select a subset $\mathbf{n}$ of those comparisons of size S that maximizes the discriminability of the output histograms. To achieve this, an iterative heuristic approach called stochastic hill climbing is adopted for obtaining an approximate solution, since the exact solution is intractable due to the combinatorial nature of the problem.

The related work is either relying on boosting algorithm for the selection of the optimal LBP features or iterative method for solving optimization with heavy computational cost. In this work, however, we propose an inexpensive, closed-form solution for learning weight-optimal local binary patterns (WoLBP), which can be easily extended to many LBP variants and should lead to performance boost. For this very reason, we only benchmark our proposed WoLBP against traditional LBP implementation.

2 Weight-optimal Local Binary Patterns

In this section, we will first review the formulation of traditional local binary patterns and then detail the formulation of the proposed weight-optimal local binary patterns.

2.1 Traditional Local Binary Patterns

We start by formulating the traditional LBP operator first introduced by Ojala et al. [19]. The basic idea of this approach is demonstrated in Figure 1. Here we have shown both the 3×3 patch and 5×5 patch. 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 to characterize the local texture.

The LBP texture for center point (x_c, y_c) can be represented as:

$$LBP(x_c, y_c) = \sum_{n=0}^{L-1} s(i_n - i_c)2^n \tag{1}$$

where i_n denotes the intensity of the n^{th} 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.

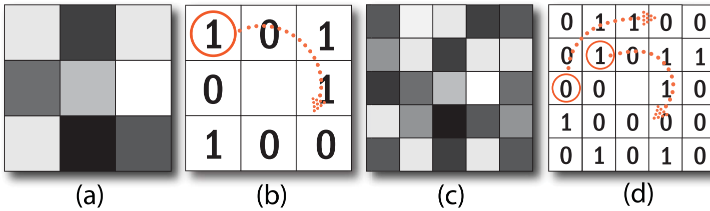

Fig. 1. (a) 3×3 neighborhood, (b) LBP encoding, the 8-bit representation for the center pixel is 10110010 and 178 in decimal, (c) 5×5 neighborhood, and (d) LBP encoding using both radii around the center pixel.

During the formulation of the LBP feature, there are many knobs one can play with and result in totally different LBPs. For example, the ordering of the bit string matters if it is converted to a decimal number, the choice of the pivot point (center point), and the choice of bases. More discussions can be found in [13].

Varying Base One can vary the base used for forming the decimal number from the bit string. Instead of using base 2 for conversion as is universally adopted [21], fractional bases (e.g., 1.6, 0.76) or other integer bases (e.g., 3, 4) can also be used. Unleashing the restriction of using only base 2 for decimal conversion, much more diversity can be achieved when encoding LBPs.

Varying Pivot/Center In the case of 3×3 neighborhood, the center pixel for thresholding neighbors is usually the physical center of the neighborhood. However, one can vary the center in a larger neighborhood as shown in Figure 2. Each pivot (thresholding center) gives different bit string, so varying the center will also provide much more diversity.

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. One can easily vary the ordering of the neighbors as shown in Figure 2, and thus lead to different formulation of the LBPs.

2.2 Weight-optimal Local Binary Patterns

All the possible variations mentioned above can be determined empirically, for instance, the choice of center point, the base, and the ordering of the neighbors. In this work, we propose to reformulate the problem of LBP encoding by using a learning framework for obtaining the optimal weights.

First, we need to reformulate the LBP encoding problem into matrix multiplication. The traditional way of encoding LBP feature is to use a 3×3 window to scan through the entire image. At each 3×3 patch, perform the encoding

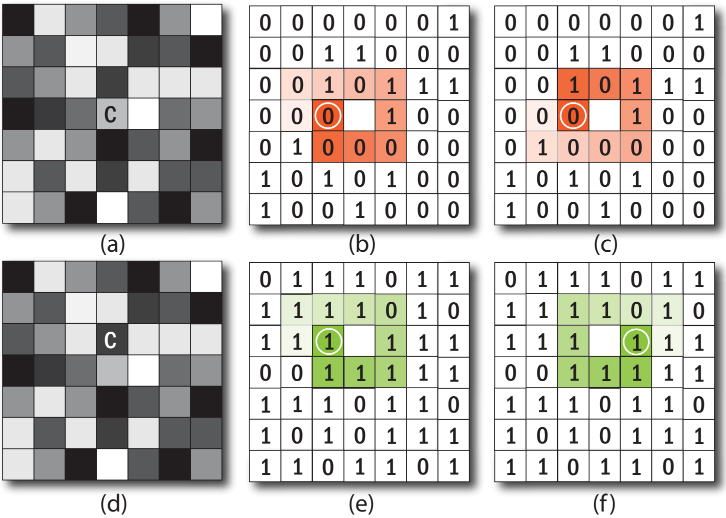

Fig. 2. Example of the LBP encoding scheme [13]: (a) 7×7 neighborhood with the thresholding center denoted as C , (b,c) 2 ordering schemes of (a), (d) 7×7 neighborhood with different thresholding center, and (e,f) 2 possible ordering schemes of (d). Different choices of the thresholding center and ordering schemes result in different LBP code.

using Equation 1. However, such formulation is neither efficient, nor provides insight towards an optimal weight learning scheme.

Instead of scanning through the entire image using small window and compare the neighborhood values to its center point, a simple convolution of the image with 8 difference masks, followed by simple binarization can achieve the same goal. As shown in Figure 3, we can use 8 difference masks of size 3×3 to convolve with the face image. The 8 resulting bitmaps are shown around the original face image. The traditional LBP is simply a weighted sum of all the bitmaps using the weight vector $\mathbf{w} = [2^7, 2^6, 2^5, 2^4, 2^3, 2^2, 2^1, 2^0]$. Therefore, the reformulation of the LBP can be shown as:

$$\mathbf{y} = \sum_{i=1}^{8} \sigma(\mathbf{h}_i * \mathbf{f}) \cdot \mathbf{w}_i \tag{2}$$

where $\mathbf{f} \in \mathbb{R}^d$ is the original image, $\mathbf{h}_i$ are the difference masks, σ is the binarization operator, and $\mathbf{y} \in \mathbb{R}^d$ is the resulting LBP image. Note that only the binarization is non-linear operation.

Now, we are one step closer towards the formulation of the WoLBP. In the context of N training images from K classes, we re-arrange them in the following way: the N training images are vectorized and becomes one column in the data matrix $\mathbf{F} \in \mathbb{R}^{d \times N}$, and for each image in $\mathbf{F}$ we apply convolutional mask $\mathbf{h}_1$ to obtain $\mathbf{X}_1 \in \mathbb{R}^{d \times N}$. Then we repeat for $\mathbf{h}_2$ to $\mathbf{h}_8$ to obtain $\mathbf{X}_2$

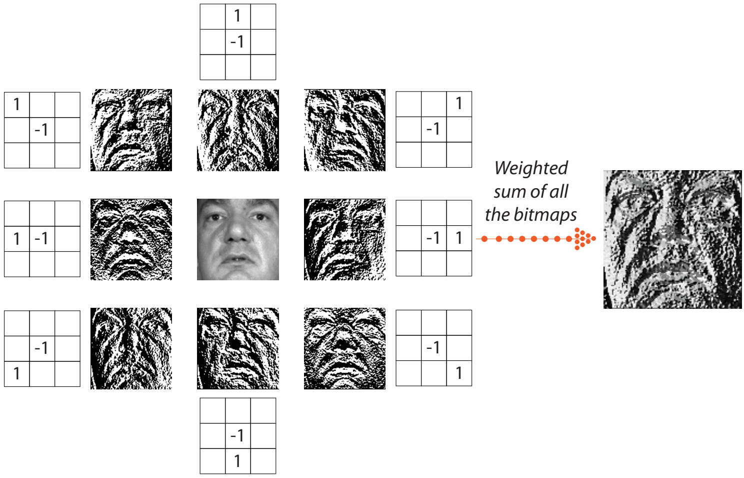

Fig. 3. Reformulation of the traditional LBP encoding using convolution.

to $\mathbf{X}_8$. Stacking all $\mathbf{X}_i$'s would give us the new bitmap matrix $\mathbf{X} \in \mathbb{R}^{8d \times N}$. The weight vector $\mathbf{w}$ is now re-written as weight matrix $\boldsymbol{\Omega} \in \mathbb{R}^{d \times 8d}$, where $\boldsymbol{\Omega} = [\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{\Omega}_3, \boldsymbol{\Omega}_4, \boldsymbol{\Omega}_5, \boldsymbol{\Omega}_6, \boldsymbol{\Omega}_7, \boldsymbol{\Omega}_8]$, and $\boldsymbol{\Omega}_i = \mathbf{w}_i \cdot \mathbf{I}$. In this way, the LBP image for all the N training images can be found in $\mathbf{Y} \in \mathbb{R}^{d \times N}$ using:

$$\mathbf{Y} = \boldsymbol{\Omega}\mathbf{X} \tag{3}$$

As shown in Figure 4, the weight matrix of the traditional LBP is a horizontal stacking of 8 diagonal matrices, each is a multiple of the identity matrix, and the multiple is defined by the weight vector $\mathbf{w}$. One generalization is as follows. For each of the 8 diagonal weight matrices $\boldsymbol{\Omega}_i$, we allow the diagonal to take d different values corresponding to the d dimensions of each bitmap. A even further generalization is allowing $\boldsymbol{\Omega}$ to be a full matrix, as shown in Figure 5, and when multiplied with the bitmap matrix $\mathbf{X}$, the generated LBP image can be somewhat optimal.

Here, the objective of the optimization is to make the LBP images $\mathbf{Y}$ have the best class separation, and thus lead to better classification performance. Fisher ratio is one way to characterize the class separability by simultaneously maximizing the between-class scatter and minimizing the within-class scatter. Note that the only non-linear part within the LBP formulation, binarization, has been taken care of by stacking all the bitmaps in the matrix $\mathbf{X}$, and a linear method is sufficient to learn an optimal weight matrix $\boldsymbol{\Omega}$.

So we are trying to solve for the following optimization:

$$\text{maximize} \frac{|\mathbf{S}_b^{\mathbf{Y}}|}{|\mathbf{S}_w^{\mathbf{Y}}|} = \underset{\boldsymbol{\Omega}}{\text{maximize}} \frac{|\boldsymbol{\Omega}\mathbf{S}_b^{\mathbf{X}}\boldsymbol{\Omega}^\top|}{|\boldsymbol{\Omega}\mathbf{S}_w^{\mathbf{X}}\boldsymbol{\Omega}^\top|} \tag{4}$$

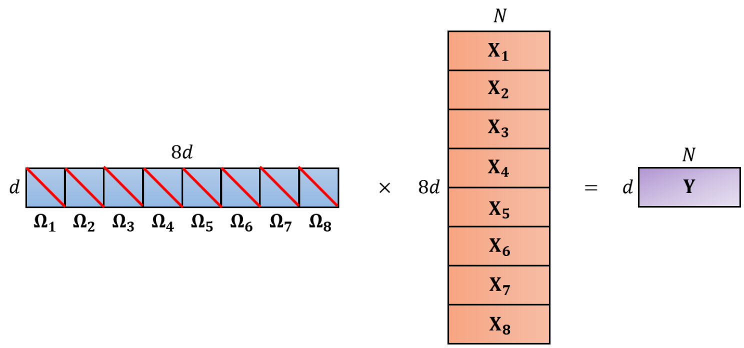

Fig. 4. Traditional LBP in matrix multiplication form.

whose optimality can be found by solving the eigenvalue problem:

$$\mathbf{\Omega}\mathbf{\Lambda} = \left((\mathbf{S}_w^{\mathbf{X}})^{-1}\mathbf{S}_b^{\mathbf{X}}\right)\mathbf{\Omega} \tag{5}$$

where

$$\mathbf{S}_w^{\mathbf{X}} = \sum_{i=1}^{K}\sum_{j\in C_i}(\mathbf{x}_j - \mu_i)(\mathbf{x}_j - \mu_i)^\top \tag{6}$$

$$\mathbf{S}_b^{\mathbf{X}} = \sum_{i=1}^{K}(\mu_i - \mu)(\mu_i - \mu)^\top \tag{7}$$

where μ_i are the mean vector of all the $\mathbf{x}_i$'s belonging to class i (denoted as C_i), and μ is the global mean vector. $\mathbf{x}_1 \ldots \mathbf{x}_N$ are the columns of bitmap matrix $\mathbf{X}$.

Solving Equation 5 would give the optimal weight matrix $\mathbf{\Omega}$ which leads to the highest Fisher ratio for the LBP image matrix $\mathbf{Y}$. The optimal weight matrix can be seen as a linear transformation matrix that reduces the dimensionality from $8d$ to d. Please note that this WoLBP learning procedure is different from regular Linear Discriminant Analysis (LDA) because in LDA, a transformation matrix $\mathbf{W}$ is learned to reduce the dimensionality of $\mathbf{Y}$ from d to d' where $d' < d$. Whereas in WoLBP procedure, the learning is restricted to feature encoding which maps the dimension from $8d$ to d on the bitmap matrix $\mathbf{X}$. In short, we have carried out feature encoding learning in WoLBP, not subspace learning for images.

3 Experiments

In this section, the effectiveness of the proposed WoLBP is validated in the context of face recognition. We detail the database used in the experiments first, and then the experimental setup and results.

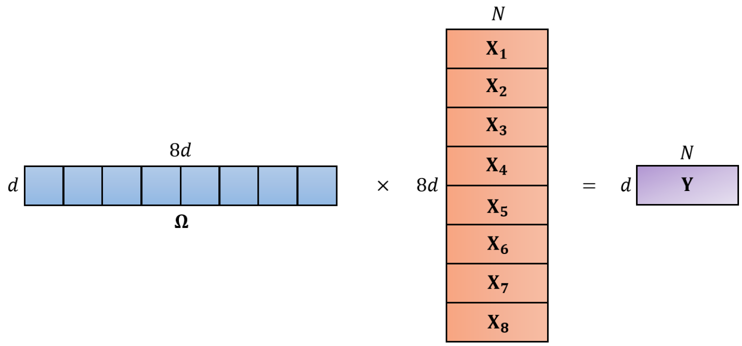

Fig. 5. Allowing $\mathbf{\Omega}$ to be a full matrix and form the WoLBP.

3.1 Database

In this work, we utilize the largest frontal face database that is publicly available: NIST's Face Recognition Grand Challenge (FRGC) ver2.0 database [20] to validate the effectiveness of our proposed method, which is also adopted in [17].

The FRGC database 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, as shown in Figure 6.

Fig. 6. Examples from the FRGC ver2.0 database: (a1,a2) controlled and uncontrolled still, (b1,b2) cropped full face.

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. **FRGC Experiment 1** is the largest experiment in the FRGC protocol which involves over 256 million face matching comparisons.

One of the latest trends in face recognition community seems to be working on unconstrained dataset such as the LFW [2], 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 [16]. 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.

3.2 Experimental Setup

In our experiments, we follow NIST's **FRGC Experiment 1** protocol which involves 1-to-1 matching of 16,028 target images to themselves (~256 million pair-wise face match comparisons). The WoLBP training is carried out on the generic training set. After obtaining the optimal weight matrix $\mathbf{\Omega}$, it is applied on all the images in the target set. In this experiment, we do not resort to any subspace learning algorithms. The normalized cosine distance (NCD) measurement is adopted to compute similarity matrix between target set images:

$$d(\mathbf{x}, \mathbf{y}) = \frac{-\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}. \tag{8}$$

Compared to other commonly used distance measurement such as ℓ_1-norm, ℓ_2-norm, and the Mahalanobis distance, NCD exhibits the best result.

The result of each algorithm is a similarity matrix with the size of $16,028 \times 16,028$ whose entry SimM_{ij} is the NCD between the feature vector of target image i and target image j. The performance of WoLBP and traditional LBP is analyzed using verification rate (VR) at 0.1% (0.001) false accept rate (FAR), equal error rate (EER), and receiver operating characteristic (ROC) curves.

3.3 Experimental Results

The VR at 0.1% FAR and EER are shown in Table 1. ROC curves are shown in Figure 7 for 32×32 and 64×64 image size respectively. As can be seen, the WoLBP performs significantly higher than the traditional LBP which has hard-coded encoding scheme. We have also found the same trend on other frontal face databases such as CMU Multi-PIE and YaleB+ database. However, the scale of these databases are no comparison with the FRGC ver2.0 database which involves more than 256 million face matches. Therefore, we do not report the results and ROC curves for those databases for the sake of brevity.

Table 1. VR at 0.1% FAR and EER for the FRGC evaluation.

	32×32		64×64	
	VR at 0.1% FAR	***EER***	***VR at 0.1% FAR***	***EER***
WoLBP	0.807	0.040	0.801	0.042
LBP	0.516	0.131	0.496	0.137
Pixel	0.349	0.170	0.350	0.167

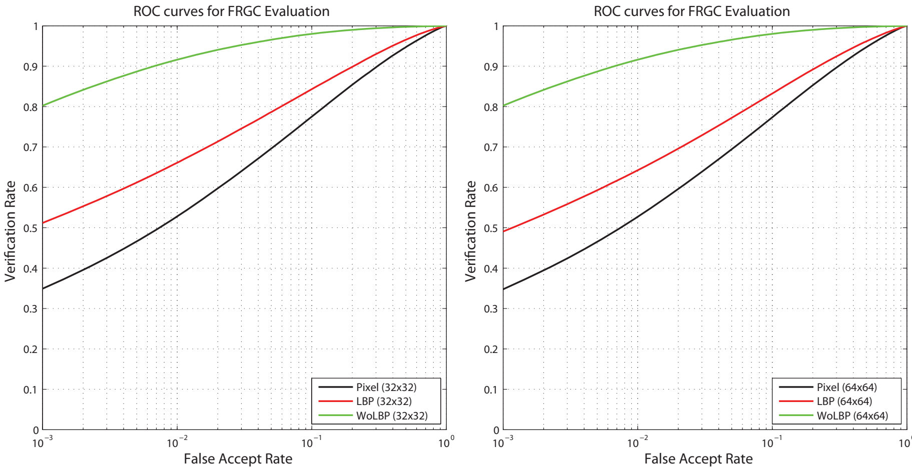

Fig. 7. ROC curves on FRGC face images of size 32×32 (left) and 64×64 (right).

3.4 Discussions

The optimality discussed in this work is solely determined by Fisher ratio. Of course, there can be other optimally obtained local binary patterns via other criteria. It is also worth noted that in order to make Fisher ratio work properly, the homoscedasticity property has to hold, meaning the training data from different classes should all be uni-modal Gaussian distributed with equal covariance. For natural images, this is most likely true. However, readers are encouraged to check the homoscedasticity property when applying the WoLBP technique discussed in this work.

4 Conclusions

In this work, we have proposed a learning paradigm for obtaining weight-optimal local binary patterns (WoLBP). We first re-formulate the LBP problem into matrix multiplication with all the bitmaps flattened and then resort to the Fisher ratio criterion for obtaining the optimal weight matrix for LBP encoding. The solution is closed form and can be easily solved using one eigen-decomposition. The experimental results have shown that the WoLBP gains significant performance improvement over traditional LBP, and such WoLBP learning procedure can be directly ported to many other LBP variants to further improve their performances.

References

1. Ahonen, T., Hadid, A., Pietikänen, M.: Face recognition with local binary patterns. In: European Conference on Computer Vision. Lecture Notes in Computer Science. pp. 469–481 (2004)
2. Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Tech. Rep. 07-49, Univ. of Massachusetts, Amherst (oct 2007)
3. Jin, H., Liu, Q., Lu, H., Tong, X.: Face detection using improved lbp under bayesian framework. In: Image and Graphics, Proc. 3rd Int'l Conf. on. pp. 306–309 (dec 2004)
4. Juefei-Xu, F., Cha, M., Heyman, J.L., Venugopalan, S., Abiantun, R., Savvides, M.: Robust local binary pattern feature sets for periocular biometric identification. In: Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on. pp. 1–8. IEEE (2010)
5. Juefei-Xu, F., Cha, M., Savvides, M., Bedros, S., Trojanova, J.: Robust periocular biometric recognition using multi-level fusion of various local feature extraction techniques. In: IEEE 17th International Conference on Digital Signal Processing (DSP) (2011)
6. Juefei-Xu, F., Luu, K., Savvides, M., Bui, T.D., Suen, C.Y.: Investigating age invariant face recognition based on periocular biometrics. In: Biometrics (IJCB), 2011 International Joint Conference on. pp. 1–7. IEEE (2011)
7. Juefei-Xu, F., Pal, D.K., Savvides, M.: Hallucinating the full face from the periocular region via dimensionally weighted k-svd. In: IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW). IEEE (2014)
8. Juefei-Xu, F., Savvides, M.: Can your eyebrows tell me who you are? In: Signal Processing and Communication Systems (ICSPCS), 2011 5th International Conference on. pp. 1–8. IEEE (2011)
9. Juefei-Xu, F., Savvides, M.: Unconstrained periocular biometric acquisition and recognition using cots ptz camera for uncooperative and non-cooperative subjects. In: Applications of Computer Vision (WACV), 2012 IEEE Workshop on. pp. 201–208. IEEE (2012)
10. Juefei-Xu, F., Savvides, M.: An augmented linear discriminant analysis approach for identifying identical twins with the aid of facial asymmetry features. In: IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW) (2013)
11. Juefei-Xu, F., Savvides, M.: An image statistics approach towards efficient and robust refinement for landmarks on facial boundary. In: Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on. pp. 1–8. IEEE (2013)
12. Juefei-Xu, F., Savvides, M.: Facial ethnic appearance synthesis. In: European Conference on Computer Vision (ECCV) Workshops. Springer (2014)
13. Juefei-Xu, F., Savvides, M.: Subspace based discrete transform encoded local binary patterns representations for robust periocular matching on nists face recognition grand challenge. IEEE Transactions on Image Processing (2014)
14. Lei, Z., Pietikainen, M., Li, S.: Learning discriminant face descriptor. Pattern Analysis and Machine Intelligence, IEEE Transactions on 36(2), 289–302 (Feb 2014)
15. Lei, Z., Yi, D., Li, S.: Discriminant image filter learning for face recognition with local binary pattern like representation. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. pp. 2512–2517 (June 2012)

16. Leibo, J.Z., Liao, Q., Poggio, T.: Subtasks of unconstrained face recognition. International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. (VISAPP) (2014)
17. Liao, S., Zhu, X., Lei, Z., Zhang, L., Li, S.Z.: Learning multi-scale block local binary patterns for face recognition. In: International Conference on Biometrics (ICB). pp. 828–837 (2007)
18. Maturana, D., Mery, D., Soto, A.: Learning discriminative local binary patterns for face recognition. In: Automatic Face Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on. pp. 470–475 (March 2011)
19. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. Pattern Recognition 29(1), 51–59 (1996)
20. Phillips, P.J., Flynn, P.J., Scruggs, T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Min, J., Worek, W.: Overview of the face recognition grand challenge. In: Computer Vision and Pattern Recognition. CVPR. IEEE Computer Society Conf. on. vol. 1, pp. 947–954 (jun 2005)
21. Pietikäinen, M., Hadid, A., Zhao, G., Ahonen, T.: Computer Vision Using Local Binary Patterns. Springer (2011)
22. Shan, C.: Learning local binary patterns for gender classification on real-world face images. Pattern Recognition Letters 33, 431–437 (2012)
23. Shan, C., Gong, S., McOwan, P.W.: Robust facial expression recognition using local binary patterns. In: Image Processing. ICIP. IEEE Int'l Conf. on. vol. 2, pp. 370–376 (sep 2005)
24. Sun, N., Zheng, W., Sun, C., Zou, C., Zhao, L.: Gender classification based on boosting local binary pattern. In: Proc. 3rd Int'l Symposium on Neural Networks (ISNN). Lecture Notes in Computer Science, vol. 3972, pp. 194–201. Springer (2006)
25. Sun, Z., Tan, T., Qiu, X.: Graph matching iris image blocks with local binary pattern. In: Proc. Advances in Biometrics, Int'l Conf. (ICB), Lecture Notes in Computer Science 3832, Springer. pp. 366–373 (2006)
26. Zhang, H., Zhao, D.: Spatial histogram features for face detection in color images. In: Proc. Advances in Multimedia Information Processing, 5th Pacific Rim Conf. on Multimedia (PCM). vol. 1, pp. 377–384 (2004)