Discriminant Image Filter Learning for Face Recognition with Local Binary Pattern Like Representation

Zhen Lei Dong Yi Stan Z. Li*
Center for Biometrics and Security Research & National Laboratory of Pattern Recognition,
Institute of Automation, Chinese Academy of Sciences,
95 Zhongguancun Donglu, Beijing 100190, China.
{zlei,dyi,szli}@cbsr.ia.ac.cn

Abstract

Local binary pattern (LBP) and its variants are effective descriptors for face recognition. The traditional LBP like features are extracted based on the original pixel or patch values of images. In this paper, we propose to learn the discriminative image filter to improve the discriminant power of the LBP like feature. The basic idea is after the image filtering with the learned filter, the difference of pixel difference vectors (PDVs) between the images from the same person is consistent and the difference between the images from different persons is enlarged. In this way, the LBP like features extracted from the filtered images are considered to be more discriminant than those extracted from the original images. Moreover, a coupled discriminant image filters learning method is proposed to deal with the heterogenous face images matching problem by reducing the feature gap between the heterogeneous images. Experiments on FERET, FRGC and a VIS-NIR heterogeneous face databases validate the effectiveness of our proposed image filter learning method combined with LBP like features.

1. Introduction

Face recognition has attracted much attention due to its potential value for applications and its theoretical challenges. In real world, the face images are usually affected by different expressions, poses, occlusions and illuminations, and the difference of face images from the same person could be larger than that from different ones. Therefore, how to extract robust and discriminant features which make the intra-person faces compact and enlarge the margin among different persons has become a critical and difficult problem in face recognition.

Up to now, many face representation approaches have been introduced, including subspace based holistic features and local appearance features [14]. Typical holistic features include the well known Principal Component Analysis (PCA) [25], Linear Discriminant Analysis (LDA) [2], Independent Component Analysis (ICA) [4] etc.

Local appearance features, as opposed to holistic features like PCA and LDA, have certain advantages. They are more stable to local changes such as illumination, expression and inaccurate alignment. Gabor [17, 12] and local binary patterns (LBP) [1] are two representative features. Gabor wavelets capture the local structure corresponding to specific spatial frequency (scale), spatial locality, and selective orientation. It has been demonstrated to be discriminative and robust to illumination and expression changes. Local binary patterns (LBP) which describes the neighboring changes around the central point, is a simple yet effective way to represent faces. It is invariant to monotone transformation and is robust to illumination changes to some extent. The combination of Gabor and LBP further improves the face recognition performance. A lot of work has been proposed in this branch [28, 27, 13].

1.1. Related Work

Figure 1. Three-step way to extract LBP like feature.

In general, LBP like feature extraction can be decomposed into three steps (Fig. 1). First, an image filter is applied to reduce the noise affection and enhance the useful information. Second, certain pixel patterns on the filtered image are sampled and compared. Third, the encoded image is derived based on the pixel comparison results and encoding rules. From this view, in original LBP [1], the first

*Corresponding author.

978-1-4673-1228-8/12/$31.00 ©2012 IEEE

filtering step is skipped and the LBP feature is extracted from the original image directly. The neighboring pixel values are compared with the central point and the LBP feature is encoded with a uniform pattern definition. In LGBP [28], a bank of Gabor filters with different scales and orientations are first applied and the LGBP feature is extracted from the Gabor magnitude responses. Recently, there is some work to learn the LBP encoder at the third step. Cao et al. [3] utilize unsupervised methods (random-projection tree and PCA tree) to learn the encoder and the PCA dimension reduction method is applied to get a compact face descriptor. Guo et al. [7] propose a supervised learning approach based on Fisher separation criterion to learn the encoder of LBP. In [19], authors propose to construct a decision tree for each region to encode the pixel comparison result and in [20], a heuristic algorithm is used to find the optimal pixel comparison pairs for discriminative face representation. To the best of our knowledge, there is little work on image filter learning for LBP like feature extraction. The most related work is the "Volterrafaces" [10] in which various "Volterra" kernels are learned and a vote mechanism is adopted for face recognition. However, it is not relevant to LBP like feature extraction.

1.2. Our Contribution

In this work, our attention is focused on the first step of LBP like feature extraction and the last two steps adopt the same way as in the original method. For most existing LBP related methods, the image filter in the first step is defined in an *ad hoc* way and it is difficult to obtain its optimal formulation. In this paper, we propose to learn the image filter in a data-driven way. A discriminant analysis is applied to learn the optimal filter so that after filtering, more discriminative information useful for LBP like feature can be explored. The LBP like features extracted from these responses are hence expected to be more effective for face representation.

The second contribution of this paper is to propose a coupled discriminant image filters learning method for heterogeneous image matching. By appropriate formulation, coupled image filters for different modalities are learned simultaneously so that the LBP like features extracted from the filtered images are similar. The feature difference of heterogeneous images is reduced and it makes the matching easier.

2. Discriminative Image Filter Learning for Local Binary Pattern

Without loss of generality, we take multi-scale block local binary pattern (MBLBP) [16] for example to show how we can learn the image filters to improve the discriminative power of LBP codes. In MBLBP, the mean values of local regions rather than the pixel values are compared. From the three-step view (Fig. 1), the mean filter is first applied to the image and then the ordinary LBP codes are extracted. The mean filter is able to reduce the noise to some extent, however, it may not be the optimal one to enhance the discriminative ability for face recognition.

In this work, we propose to learn an image filter that explores discriminative information for consequent face representation and recognition. That is, to reduce the variances of intra persons and meanwhile enlarge the margin between images from different persons. Given an image I, its filtered image is denoted as $f(I)$. Considering the LBP pattern sampling strategy, the pixel difference vector (PDV) can be grouped as $df(I)^p = [f(I)^{p_1} - f(I)^p, f(I)^{p_2} - f(I)^p, \cdots, f(I)^{p_d} - f(I)^p]$, where $f(I)^{p_i}$ is the pixel value of filtered image at position p_i, $\{p_1, p_2, \cdots, p_d\} \in Neighbor(p)$ and d is the number of neighbors. Note that $df(I)^p$ is a row vector. The purpose of image filter learning is to find the filter f so that after the image filtering, the PDVs of images from the same person are similar. Here, we use the Fisher criterion [6] to evaluate the discriminative ability of the pixel difference vector. Let $df(I)^p_{ij}$ be the p-th PDV of j-th sample from class i, and $df(I)_{ij} = [df(I)^1_{ij}, df(I)^2_{ij}, \cdots, df(I)^N_{ij}]$ be the PDV set from the j-th image of class i, where N is the PDV number for each image, the within and between class scatters can then be computed as

$$
\begin{aligned}
S_w &= \sum_{i=1}^{L}\sum_{j=1}^{C_i}(df(I)_{ij} - df(m)_i)(df(I)_{ij} - df(m)_i)^T \\
S_b &= \sum_{i=1}^{L} C_i(df(m)_i - df(m))(df(m)_i - df(m))^T
\end{aligned} \tag{1}
$$

where L is the total class number and C_i is the sample number of i-th class. $df(m)^p_i$ is the mean vector of PDVs at position p on filtered images from the i-th class and $df(m)^p$ is the total mean vector of PDVs at position p over the sample set. $df(m)_i = [df(m)^1_i, df(m)^2_i, \cdots, df(m)^N_i]$ and $df(m) = [df(m)^1, df(m)^2, \cdots, df(m)^N]$ are augmented vectors by concatenating mean vectors over different positions.

Under linear assumption, suppose the image filter vector to be w, and the value of filtered image at position p can be represented as $f(I)^p = w^T I^p$, where I^p denotes the image patch vector centered at position p. Similarly, the PDV $df(I)^p$ can be represented as $df(I)^p = w^T dI^p$. Substituting it into Eq. 1, we get

$$
\begin{aligned}
S_w &= w^T(\sum_{i=1}^{L}\sum_{j=1}^{C_i}(dI_{ij} - dm_i)(dI_{ij} - dm_i)^T)w \\
&= w^T \hat{S_w} w \\
S_b &= w^T(\sum_{i=1}^{L} C_i(dm_i - dm)(dm_i - dm)^T)w \\
&= w^T \hat{S}_b w
\end{aligned} \tag{2}
$$

where dI_{ij} is PDVs extracted from the j-th image of class i, dm_i is the mean PDVs matrix for the i-th class and dm is the total mean PDVs matrix.

The larger the ratio of S_b to S_w, the more discriminant information explored by the image filter. The optimal solution w which maximizes the ratio of S_b to S_w can be obtained by solving the generalized eigenvalue problem $\hat{S_b}w = \lambda\hat{S_w}w$ with its leading eigenvalue.

3. Coupled Image Filters Learning for Heterogeneous Images

Recently, more and more attention has been paid to heterogeneous face image matching (multi-modal face recognition) problem. Heterogeneous faces (faces with different modalities) are defined as faces which are captured in different environments or by different devices, *e.g.*, visual (VIS) vs. near infrared (NIR), VIS vs. Sketch etc, which are common in many real applications like law enforcement and video surveillance. In heterogeneous face image matching cases, the image appearance differs so much that the traditional texture based face recognition methods don't work well on them. Previous work mainly focuses on transforming the heterogeneous face images into the same modality for matching or developing an advanced classifier that is robust to the modality gap of features extracted from heterogeneous images. In this paper, we try to reduce the gap from different modalities at the feature level to simplify the heterogeneous face recognition problem. From the three-step view (Fig. 1), Zhang et al. [29] has proposed a coupled encoding method at the third step to reduce the difference of heterogeneous features. Our work focuses on the first step. Coupled image filters are learned so that the responses of filters are as similar as possible for heterogeneous images from the same person, and the features extracted from the filtered heterogeneous images are supposed to be more consistent with each other. In the following, we take LBP like feature as an example to show how the coupled discriminative image filters can be learned and applied to face representation.

As mentioned before, the LBP like feature is encoded on a series of difference of pixel pairs, which are called pixel difference vectors (PDVs). The purpose of coupled image filters learning is to reduce the difference of PDVs for the heterogeneous images from the same person and meanwhile enlarge the difference margin of them for those images from the different subjects. Let I^V and I^M be the face images with two modalities (e.g., visual and near infrared modalities) and their filtered images are denoted as $f(I^V)$ and $f(I^M)$. Suppose $df(I^V)_{ij}^p$ and $df(I^M)_{ij}^p$ be the p-th heterogeneous PDVs of j-th sample pair from the i-th class. Following the Fisher criterion stated in Sec. 2, the objective of our coupled image filters learning can be formulated to maximize the ratio of between class scatter to within class scatter. Denoting

$$
\begin{aligned}
S_w^{VM} &= \sum_{i=1}^{L}\sum_{j=1}^{C_i}(df(I^V)_{ij} - df(m^M)_i)(df(I^V)_{ij} - df(m^M)_i)^T \\
S_b^{VM} &= \sum_{i=1}^{L} C_i(df(m^V)_i - df(m^M))(df(m^V)_i - df(m^M))^T
\end{aligned}
\tag{3}
$$

where $df(I^V)_{ij}$, $df(I^M)_{ij}$, $df(m^V)_i$, $df(m^M)_i$, $df(m^V)$, $df(m^M)$ are defined similarly as in Sec. 2 and the superscript ^{V} or ^{M} is the modality indicator. The between and within class scatters can then be defined as

$$
\begin{aligned}
S_w &= S_w^{VV} + S_w^{MM} + S_w^{VM} + S_w^{MV} \\
S_b &= S_b^{VV} + S_b^{MM} + S_b^{VM} + S_b^{MV}
\end{aligned}
\tag{4}
$$

Under linear assumption, the filtered images $f(I^V)$ and $f(I^M)$ at position p can be formulated as $f(I^{V^p}) = {w^V}^T I^{V^p}$ and $f(I^{M^p}) = {w^M}^T I^{M^p}$, where I^{V^p} and I^{M^p} are original image patch vectors centered at position p for heterogeneous image pair and w^V and w^M are coupled image filter vectors. Substituting it into Eq. 4, followed by matrix operations, we can get the form as $S_w = w^T\tilde{S_w}w$ and $S_b = w^T\tilde{S_b}w$, where $w = [w^V; w^M]$. The optimal solution w that maximizes the ratio of S_b to S_w can then be obtained by solving the generalized eigenvalue problem $\tilde{S_b}w = \lambda\tilde{S_w}w$ with its leading eigenvalue. The coupled discriminative filters w^V and w^M are finally obtained by splitting w appropriately.

4. Image Filter Learning based Descriptor

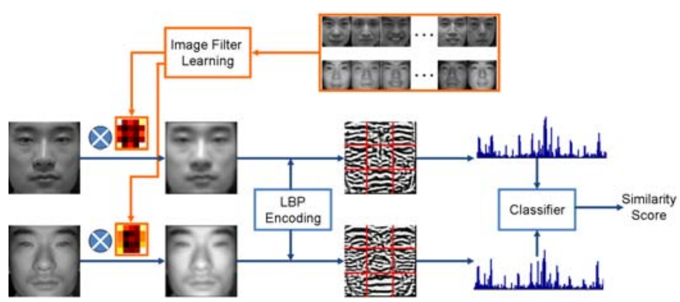

Figure 2. The pipeline of face representation and recognition with image filter learning based descriptor.

Fig. 2 illustrates the pipeline of image filter learning based method for face representation and recognition. In training phase, the (coupled) discriminative image filter is learned for homogenous or heterogenous face images. In testing phase, given two input images, they are first filtered with the learned filter(s) and then the LBP like codes are encoded from the filtered images. After that, the histogram based features are extracted to represent face images. Proper classifier is finally adopted to measure the similarity of the input two images.

Extended LBP Coding. Ordinary LBP models the relations between the central pixel and its neighboring pixels. In this work, we further model the absolute values of difference of the central point and the neighboring pixels to enhance the discriminative power of the descriptor. This method is firstly proposed by Huang et al. [8] and applied to 3D depth face recognition. They use three bits to model the difference between the central pixel and neighboring pixels since they find most of the difference values are smaller than 8. Here, we first normalize the difference values and then use one bit to model the difference for the reason of noise tolerance and efficiency of feature matching. Suppose $d = [d_0, d_1, \cdots, d_7]^T$ is the difference vector between the central point and its surrounding pixels, it is first normalized into [0,1] with its minimal value d_{min} and maximal value d_{max} as $\tilde{d}_i = (d_i - d_{min})/(d_{max} - d_{min})$. The values which are larger than 0.5 is set to 1 and those which are not larger than 0.5 is set to 0. In this way, a 8-bit binary string can be obtained and its decimal value is considered as a supplemental code to the original LBP. This improvement is denoted as the extended LBP (ELBP) in the left paper.

Classifier. For traditional homogeneous face recognition, we use the histogram intersection measure to compute the similarity of two images. For heterogenous face image matching, we use the linear coupled spectral regression (LCSR) [11] to find the common discriminant subspace and the cosine distance is adopted to compute the similarity of heterogenous image pair. The common discriminant subspace is learned from the training set.

Fusion. In order to improve the robustness of the results, we adopt multi-scale analysis by varying the radius of LBP sampling. We use the simple sum rule [9] to fuse the similarity scores of multi-scale LBPs.

5. Experiments

We compare our image filter learning based method with some state-of-the-art descriptors. For homogeneous face recognition, FERET [23] and FRGC [22] face databases are used to evaluate the performance of different methods. For heterogenous face image matching, we compare the performance of different methods on a visual image (VIS) and near infrared (NIR) face database [15].

5.1. Data Description

The FERET database is one of the largest publicly available databases. The training set contains 1002 images. In test phase, there are one gallery set containing 1196 images from 1196 subjects, and four probe sets (fb, fc, dup1 and dup2) including expression, illumination and aging variations. In this experiment, we use a subset of training set containing 540 images from 270 subjects to learn the discriminant image filter. All the images are rotated, scaled and cropped into 142×120 size according to the provided eye coordinates. No further preprocessing is adopted.

The images in FRGC v1.0.4 are used in this part. The gallery set contains 943 controlled images and the probe set contains 943 uncontrolled images, both from 275 subjects. The controlled and uncontrolled images were captured under indoor and outdoor environments respectively. All the images are normalized into 142×120 size according to the provided eye coordinates and no further preprocessing is adopted.

The VIS-NIR database was collected by CBSR for heterogeneous biometric research. There are totally 5097 images, including 2095 VIS and 3002 NIR ones from 202 persons in the database. In this experiment, we use the former 100 persons with their VIS and NIR images as training set. The left images from 102 persons form the testing set. There is no intersection of images or subjects between training and testing sets. In testing phase, the gallery set consists of VIS images and the NIR images are used as the probe ones. All the images are cropped into 128×128 size according to the automatically detected eye coordinates and the Difference-of-Gaussian (DoG) preprocessing method [24] is adopted to remove the low and high frequency noise in VIS and NIR images.

5.2. Results and Discussions

5.2.1 Homogeneous Face Recognition

We compare proposed method with popular descriptors like LBP, MBLBP, LGBP, LLGP etc. For simplicity, the image filter size in this paper is fixed to 5×5 and the sampling number is set to 8. The sampling radius r is varied from $\{3, 5, 7\}$. The discriminant image filter is learned from a training subset containing 540 images from 270 subjects. All the methods are tested following the four standard testing protocols (fb, fc, dup1, dup2).

Table 1 lists the rank-1 recognition rates of different methods on four probe sets of FERET. For ease of representation, the proposed method is denoted as IFL-LBP$_s^r$ and IFL-ELBP$_s^r$, where s is the scale size of image filter and r is the sampling radius, so as the MBLBP. In MBLBP, the block size and the sampling radius are always the same as described in original paper. From Table 1, one can see that IFL-LBP significantly outperforms MBLBP, indicating that the learned image filter does improve the discriminative ability than the mean filter. Comparing single IFL-LBP with LGBP, which is a combination of 40 Gabor filters, the best performance of the single IFL-LBP is competitive with that of LGBP. It indicates that the image filter learning based method is an effective and efficient method for face representation. However, in fc probe set with lighting variation, IFL-LBP works much worse than LGBP. It may be due to that the image filter learning based method is a data-driven one and the lack of images with various lighting conditions in training set affects the robustness of the learned image filter to lighting variation. The performance of multi-scale IFL-LBP fusion is similar to the best performance of single IFL-LBP in each probe set, but more stable than the single IFL-LBP. By introducing the extended LBP coding, IFL-ELBP improves the recognition rates by 5 percent in fc probe set compared with IFL-LBP. It verifies that the ELBP is a good and useful extention of original LBP to explore more discriminant and robust information. Overall, the image filter learning based (extended) LBP feature outperforms ordinary LBP feature and achieves comparable performance with state-of-the-art descriptors, validating the effectiveness of learning based filter for face representation. It is worth noting that the image filter learning cannot only be combined with LBP, but also with other LBP like face representations, such as DT-LBP, DLBP etc.

Table 1. Comparison results (recognition rate) of proposed method with state-of-the-art methods on FERET database.

Methods	fb	fc	dup I	dup II
LBP [1]	0.97	0.79	0.66	0.64
LGBP [28]	0.98	0.97	0.74	0.71
LVP [21]	0.97	0.70	0.66	0.50
LGT [12]	0.97	0.90	0.71	0.67
HGPP [27]	0.98	0.99	0.78	0.76
LLGP [26]	0.97	0.97	0.75	0.71
DT-LBP [19]	0.99	0.63	0.67	0.48
DLBP [20]	0.99	0.48	0.68	0.55
MBLBP_3^3	0.98	0.62	0.61	0.37
MBLBP_5^5	0.98	0.41	0.57	0.35
MBLBP_7^7	0.98	0.30	0.55	0.33
MBLBP-Fusion	0.98	0.46	0.59	0.37
IFL-LBP_5^3	0.98	0.88	0.75	0.77
IFL-LBP_5^5	0.98	0.86	0.76	0.79
IFL-LBP_5^7	0.98	0.80	0.72	0.77
IFL-LBP-Fusion	0.98	0.88	0.77	0.79
IFL-ELBP-Fusion	0.99	0.93	0.76	0.78

In order to examine the generalization of the learned image filter, we further conduct experiments on FRGC database. The image filter is learned from the FERET training set. We compare the performance of proposed IFL-LBP, IFL-ELBP with LBP and MBLBP methods. Table 2 shows that all the image filter learning based methods outperform the ordinary LBP and MBLBP, indicating the good generalization of the learned image filter. The fusion of multi-scale IFL-LBP and IFL-ELBP coding can further improve the face recognition performance, which is consistent with the results on FERET database.

Table 2. Comparison results (recognition rate) on FRGC database.

Methods	Rec. rate	Methods	Rec. rate
LBP	0.13	IFL-LBP_5^3	0.49
MBLBP_3^3	0.46	IFL-LBP_5^5	0.49
MBLBP_5^5	0.34	IFL-LBP_5^7	0.49
MBLBP_7^7	0.22	IFL-LBP-Fusion	0.59
MBLBP-Fusion	0.37	IFL-ELBP-Fusion	0.61

5.2.2 Heterogeneous Faces Matching

We test the heterogeneous face image matching on VIS-NIR face database. As in former experiment, the learned coupled image filter size is fixed to be 5×5 and the sampling number is set to 8. The sampling radius is varied from $\{3, 5, 7\}$. The proposed coupled image filter learning based methods are denoted as CIFL-LBP and CIFL-ELBP. We first compare the performance of CIFL-LBP with MBLBP without LCSR learning. After extracting the histogram feature for CIFL-LBP or MBLBP, the histogram intersection is adopted to measure the similarity of different images. Table 3 lists the recognition rates of different descriptors. The coupled image filter learning based methods significantly outperform MBLBP and LBP. It beats the best result of MBLBP by 24 percent, validating that CIFL-LBP can explore more powerful discriminant information than MBLBP in heterogeneous case. The fusion of multi-scale CIFL-LBP and the extended LBP coding further improves the recognition rates by 3 and 13 percent respectively, indicating the effectiveness of multi-scale fusion and ELBP coding.

Table 3. Recognition rates of methods with histogram intersection metric on VIS-NIR database.

Methods	Rec. rate	Methods	Rec. rate
LBP	0.36	CIFL-LBP_5^3	0.55
MBLBP_3^3	0.31	CIFL-LBP_5^5	0.49
MBLBP_5^5	0.17	CIFL-LBP_5^7	0.41
MBLBP_7^7	0.17	CIFL-LBP-Fusion	0.58
MBLBP-Fusion	0.19	CIFL-ELBP-Fusion	0.68

Next, we compare CIFL-ELBP with LBP, MBLBP, SIFT [18], HoG [5] descriptors using LCSR classifier. LCSR is trained on the training set and the cosine distance is used as the similarity measure. For MBLBP and CIF-ELBP, the fusion results of three sampling radiuses are reported. Table 4 lists the recognition rates of different descriptors and Fig. 3 illustrates the corresponding Receiver Operator Characteristic (ROC) curves. The proposed CIFL-ELBP descriptor has superior performance over LBP, MBLBP, SIFT and HoG. It achieves 95% in recognition rate and 63.4% in verification rate when false accept rate is 0.001.

Table 4. Recognition rates of different descriptors with LCSR on VIS-NIR database.

Method	LBP	MBLBP	SIFT	HoG	CIFL-ELBP
Rec. rate	0.61	0.79	0.82	0.83	0.95

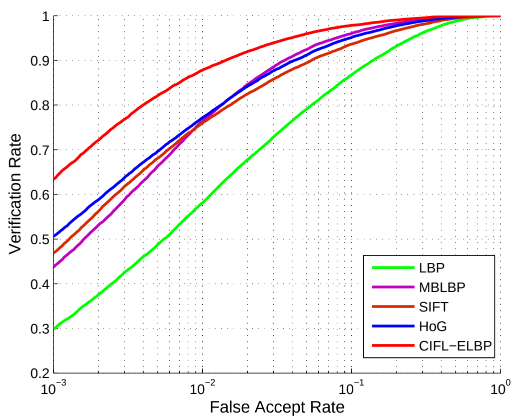

Figure 3. ROC curves of LBP, MBLBP, SIFT, HoG and CIFL-ELBP. The LCSR classifier with cosine metric is used for classification.

6. Conclusions

This paper proposes an image filter learning method for LBP like face representation. By grouping the image patches appropriately, the discriminant learning is applied to find the optimal image filter that enhances the discriminative power of LBP like feature. Moreover, we extend the image filter leaning to heterogeneous face image matching. Coupled discriminant image filters are learned simultaneously to reduce the feature gap between different face modalities. Experiments on traditional VIS face recognition and VIS-NIR heterogeneous face recognition validate the effectiveness of the proposed method. In future, we will investigate more image filter learning method combined with various face representations to improve the face recognition performance.

Acknowledgement

This work was supported by the Chinese National Natural Science Foundation Project #61070146, #61105023, #61103156, #61105037, National IoT R&D Project #2150510, European Union FP7 Project #257289 (TABULA RASA http://www.tabularasa-euproject.org), and AuthenMetric R&D Funds.

References

[1] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns:application to face recognition. *IEEE T-PAMI*, 28:2037–2041, 2006. 1, 5

[2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE T-PAMI*, 19(7):711–720, July 1997. 1

[3] Z. Cao, Q. Yin, X. Tang, and J. Sun. Face recognition with learning-based descriptor. In *CVPR*, pages 2707–2714, 2010. 2

[4] P. Comon. Independent component analysis - a new concept? *Signal Processing*, 36:287–314, 1994. 1

[5] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *CVPR*, pages 886–893, 2005. 5

[6] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification, Second Edition*. John Wiley and Sons, 2001. 2

[7] Y. Guo, G. Zhao, M. Pietikäinen, and Z. Xu. Descriptor learning based on fisher separation criterion for texture classification. In *ACCV*, pages 185–198, 2010. 2

[8] Y. Huang, Y. Wang, and T. Tan. Discriminating 3d faces by statistics of depth differences. In *ACCV*, pages 690–699, 2007. 4

[9] J. Kittler, M. Hatel, R. P. W. Duin, and J. Matas. On combining classifiers. *IEEE T-PAMI*, 20(3):226–239, January 1998. 4

[10] R. Kumar, A. Banerjee, and B. C. Vemuri. Volterrafaces: Discriminant analysis using volterra kernels. In *CVPR*, pages 150–155, 2009. 2

[11] Z. Lei and S. Z. Li. Coupled spectral regressoin for matching heterogeneous faces. In *CVPR*, 2009. 4

[12] Z. Lei, S. Z. Li, R. Chu, and X. Zhu. Face recognition with local gabor textons. In *ICB*, pages 49–57, 2007. 1, 5

[13] Z. Lei, S. Liao, M. Pietikäinen, and S. Z. Li. Face recognition by exploring information jointly in space, scale and orientation. *IEEE T-IP*, 20(1):247–256, Jan. 2011. 1

[14] S. Z. Li and A. K. Jain (ed.). *Handbook of Face Recognition*. Springer-Verlag, New York, March 2005. 1

[15] S. Z. Li, Z. Lei, and M. Ao. The hfb face database for heterogeneous face biometrics research. In *OTCBVS*, 2009. 4

[16] S. Liao, X. Zhu, Z. Lei, L. Zhang, and S. Z. Li. Learning multi-scale block local binary patterns for face recognition. In *ICB*, pages 828–837, 2007. 2

[17] C. Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE T-IP*, 11(4):467–476, 2002. 1

[18] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004. 5

[19] D. Maturana, D. Mery, and A. Soto. Face recognition with decision tree-based local binary patterns. In *ACCV*, pages 618–629, 2010. 2, 5

[20] D. Maturana, D. Mery, and A. Soto. Learning discriminative local binary patterns for face recognition. In *FG*, pages 470–475, 2011. 2, 5

[21] X. Meng, S. Shan, X. Chen, and W. Gao. Local visual primitives (lvp) for face modelling and recognition. In *ICPR*, 2006. 5

[22] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. In *CVPR*, 2005. 4

[23] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The FERET evaluation methodology for face-recognition algorithms. *IEEE T-PAMI*, 22(10):1090–1104, 2000. 4

[24] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. In *AMFG*, 2007. 4

[25] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. In *CVPR*, pages 586–591, Hawaii, June 1991. 1

[26] S. Xie, S. Shan, X. Chen, X. Meng, and W. Gao. Learned local gabor patterns for face representation and recognition. *Signal Processing*, 89(12):2333–2344, 2009. 5

[27] B. Zhang, S. Shan, X. Chen, and W. Gao. Histogram of gabor phase patterns (hgpp): A novel object representation approach for face recognition. *IEEE T-IP*, 16(1):57–68, 2007. 1, 5

[28] W. Zhang, S. Shan, W. Gao, and H. Zhang. Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition. In *ICCV*, pages 786–791, 2005. 1, 2, 5

[29] W. Zhang, X. Wang, and X. Tang. Coupled information-theoretic encoding for face photo-sketch recognition. In *CVPR*, 2011. 3