

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins



CrossMark

Extended local binary patterns for face recognition

Li Liu^{a,*}, Paul Fieguth^b, Guoying Zhao^c, Matti Pietikäinen^c, Dewen Hu^d

^a School of Information System and Management, National University of Defense Technology, 47 Yanwachi, Changsha, Hunan 410073, China

^b Department of Systems Design Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

^c The Center for Machine Vision Research, Department of Computer Science and Engineering, University of Oulu, Oulu 90014, Finland

^d School of Mechatronics and Automation, National University of Defense Technology, 47 Yanwachi, Changsha, Hunan 410073, China

ARTICLE INFO

Article history: Received 17 June 2015 Revised 28 March 2016 Accepted 9 April 2016 Available online 19 April 2016

Keywords: Face recognition Feature extraction Local binary pattern Local descriptors

ABSTRACT

This paper presents a simple and novel, yet highly effective approach for robust face recognition. Using LBP-like descriptors based on local accumulated pixel differences – Angular Differences and Radial Differences, the local differences were decomposed into complementary components of signs and magnitudes. Based on these descriptors we developed labeled dominant patterns where the most frequently occurring patterns and their labels were learned to capture discriminative textural information. Six histogram features were obtained from each given face image by concatenating spatial histograms extracted from non-overlapping subregions. A whitened PCA technique was used for dimensionality reduction to produce more compact, robust and discriminative features, which were then fused using the nearest neighbor classifier, with Euclidean distance as the similarity measure.

We evaluated the effectiveness of the proposed method on the Extended Yale B, the large-scale FERET, and CAS-PEAL-R1 databases, and found that that the proposed method impressively outperforms other well-known systems with a recognition rate of 74.6% on the CAS-PEAL-R1 lighting probe set.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Face recognition, as one of the most successful applications of image analysis and understanding, has received considerable attention in the past decades due to its challenging nature and vast range of applications [18,42]. In recent years this field progressed significantly and a number of face recognition and modeling systems have been developed [18]. Even though current face recognition systems have reached a certain level of maturity, the performance of these systems in uncontrolled environments is still far from the capability of the human vision system [18,42]. In other words, there are still many challenges associated with accurate and robust face recognition in regards to computer vision and pattern recognition, especially under unconstrained environments. Therefore, this remains a stimulating field for further research.

As a classical pattern recognition problem, face recognition primarily consists of two critical subproblems: feature extraction and classifier designation, both of which have been the subject of significant study. Generally, facial feature description plays a relatively more important role if poor features are used, and even the best classifier will fail to achieve good

* Corresponding author. Tel./fax: +86 731 84574531.

E-mail addresses: lilyliu_nudt@163.com, liuli_nudt@nudt.edu.cn (L. Liu), pfieguth@uwaterloo.ca (P. Fieguth), gyzhao@ee.oulu.fi (G. Zhao), matti. pietikainen@ee.oulu.fi (M. Pietikäinen), dwhu@nudt.edu.cn (D. Hu).

recognition results. Numerous facial feature sets were presented with excellent surveys in reports by Zhao et al. [42] and Li and Jain [18]. Nevertheless, designing useful face descriptors remains a great challenge when faced with three competing goals: computational efficiency, effective discrimination, and robustness to intra-person variations (including changes in illumination, pose, expression, age, blur, and occlusion).

Most facial extraction approaches can be categorized as utilizing either holistic features or local features. The holistic approach uses the entire face region to construct a subspace representing a face image; Influential examples includes Principle Component Analysis (PCA) [36], Linear Discriminant Analysis (LDA) [3], Locally Linear Embedding (LLE) [33] and Locally Preserving Projections (LPP) [11]. In contrast, the local features approach involves local features first being extracted from a subregion of a face and then classified by combining and comparing with corresponding local statistics. Holistic features are unable to capture local variations in face appearance, and are more sensitive to variations in illumination, expression and occlusions [18]. Local feature based approaches are advantageous in that distributions of face images in local feature space are less affected by changes in facial appearance. As a result, local feature based face recognition approaches have been widely studied in recent years.

Among local feature approaches, local binary patterns (LBP) have emerged as one of the most prominent face analysis methods since the pioneering work by Ahonen et al. [1], and have attracted increasing attention due to their distinguished advantages: (1) ease of implementation; (2) invariance to monotonic illumination changes; and (3) low computational complexity. However, the original LBP method still has multiple limitations: (i) production of long histograms; (ii) capture of only very local texture structure and not long range information; (iii) limited discriminative capability based purely on local binarized differences; and (iv) limited noise robustness. On the basis of these issues, many LBP variants [12] have been proposed to improve face recognition performance.

Our research is motivated by recent work on texture classification in [24], where four LBP-like descriptors – Center Intensity based LBP (CI-LBP), Neighborhood Intensity based LBP (NI-LBP), Radial Difference based LBP (RD-LBP) and Angular Difference based LBP (AD-LBP) – were proposed, along with multiscale joint histogram features, which were found to be highly effective to rotation invariant texture classification. We expanded this research [24] to address the face identification problem by proposing a framework and more generalized formulation of the local intensities and differences features. Specifically, the major contributions of our work are summarized as follows:

- 1. We proposed a new family of LBP-like descriptors based on local accumulated pixel differences: Angular Differences (AD) and Radial Differences (RD). The descriptors presented advantages of efficiency, complementarity to LBP, robustness, and the encoding of both microstructures and macrostructures, resulting in a more complete image representation. The extraction of the proposed descriptors did not require any training, and thus this approach showed better generalizability compared with popular learning methods, whose performance is degraded if the distribution of the testing sample varies significantly from that of the training set.
- Using the proposed descriptors, we found that the properties of the original uniform patterns introduced by Ojala et al.
 [30] did not hold true, and therefore we suggest the use of full patterns. We also proposed a labeled dominant pattern (LDP) scheme, which learns a set of dominant patterns (the most frequently occurring patterns from a set of training images) to capture discriminative textural information, but with lower dimensionality than the full pattern approach.
- 3. Extensive experiments were conducted on the Extended Yale B, the FERET and the large-scale CAS-PEAL-R1 databases. Our approach proved to be highly robust to illumination changes, as evident by our recognition rate of **74.6**% on the CAS-PEAL-R1 lighting probe set. This is, to the best of our knowledge, the highest score yet achieved on this data set.

A preliminary version of this work appeared in [22].

2. Related work

The primary motivation of this work is to design novel LBP-like descriptors and apply them to face recognition, thus the related literature focuses on LBP and its variants for face recognition.

A brief review of LBP

LBP was originally proposed for texture analysis [29,30], and was later introduced to face recognition by Ahonen et al. [1]. Since then, LBP-based facial image analysis has been one of its most popular and successful applications [6,12,17].

LBP characterizes the spatial structure of a local image texture by thresholding a 3×3 square neighborhood with the value of the center pixel and considering only the sign information to form a local binary pattern. A more general formulation defined on circular symmetric neighborhood systems was proposed by Ojala et al. [30] that allowed for multi-resolution analysis and rotation invariance. Formally, given a pixel $x_{0,0}$ in the image, the LBP pattern was computed by comparing its value with those of its *p* neighboring pixels

$$\underline{\mathbf{x}}_{r,p} = [x_{r,p,0}, \ldots, x_{r,p,p-1}]^{I}$$

that were evenly distributed in angle on a circle of radius r centered on $x_{0,0}$, as shown in Fig. 1, such that the LBP response was calculated as



Fig. 1. A central pixel and its p circularly and evenly spaced neighbors on a circle of radius r.



Fig. 2. Two commonly used neighborhood structures: (a) concentric rings of interpolated pixels, (b) a ring of image patches.

$$LBP_{r,p} = \sum_{n=0}^{p-1} s(x_{r,p,n} - x_c) 2^n,$$
(1)

where $s(\cdot)$ was the sign function, defined as follows:

$$s(x) = \begin{cases} 1 & x \ge 0\\ 0 & x < 0 \end{cases}$$
(2)

Relative to the origin at (0, 0), the coordinates of the neighbors were given by $-r\sin(2\pi n/p), r\cos(2\pi n/p)$. The intensity values of neighbors not lying on a pixel location were estimated by interpolation.

The original descriptor LBP_{r, p} produces 2^p different output values, corresponding to the 2^p different binary patterns formed by the *p* pixels in the neighborhood. Due to the rapid increase in the number of patterns for basic LBP_{r, p} as *p* is increased, it is difficult to apply LBP to large neighborhoods, therefore limiting its applicability.

In order to reduce the dimensionality of the LBP histogram feature and increase its robustness to noise, Ojala et al. [30] introduced uniform patterns, $LBP_{r,p}^{u2}$, where a binary pattern is uniform if it contains a maximum of two bitwise transitions from 0 to 1, or vice versa when the bit pattern is considered circularly. There are p(p-1) + 2 uniform patterns, and all of the remaining non-uniform patterns are accumulated into a single bin, resulting in a p(p-1) + 3 dimensional representation, and a great reduction to the 2^p dimensionality of LBP. The success of the $LBP_{r,p}^{u2}$ operator came from the experimental observations that the uniform patterns appeared to be fundamental properties of face images [1], representing salient local texture structure.

As aforementioned, the original LBP method has limitations to be addressed, and therefore, a number of extensions and modifications to it have been developed [12].

Changing neighborhood topology

Some LBP variants were proposed by changing the topology of the neighborhood from which the LBP features were computed, such as the Multi-scale Block LBP (MBLBP) [20], Three Patch LBP (TPLBP) [38] and Four Patch LBP (FPLBP) [38]. In MBLBP, the neighborhood shown in Fig. 2(b) was adopted: a $w \times w$ patch centered on the pixel of interest and n additional patches distributed uniformly in a ring of radius r around it. The LBP computation was done based on comparison of the average values of the neighboring patches against that of the center patch. MBLBP was claimed to be more robust than LBP since it encoded not only microstructures but also macrostructures of image patterns. A similar neighborhood was used by TPLBP [38], which computes patch differences, rather than single pixel or averaged pixel differences, as in LBP and MBLBP, respectively. TPLBP captured information complementary to pixel based descriptors. The Patterns of Oriented Edge

Magnitudes (POEM) descriptor proposed by Vu et al. [37] also used the same neighborhood structure as MBLBP and TPLBP, but computed LBP-like features on histograms of gradients.

Increasing discriminative power

Researchers found it helpful to preprocess an image prior to LBP computation, where the most popular preprocessing method is Gabor filtering, because Gabor and LBP provides complementary information: LBP captures local information, while Gabor filters are non-local. Two such representative works are Local Gabor Binary Pattern Histogram Sequence (LGBPHS) [41] and Histogram of Gabor Phase Patterns (HGPP) [40]. In both methods, an image is first convolved with 40 Gabor filters at five different scales and eight orientations, which is computationally expensive. In LGBPHS, the LBP method is then applied to all 40 Gabor magnitude images. In HGPP, the Gabor phase information is encoded for 90 images of the same size as the original. Due to richer information from this additional Gabor filtering stage, LGBPHS and HGPP improve face recognition performance when compared to the original LBP method. However, the Gabor filtering step is accompanied by a heavy computational burden, and the extremely high dimensionality of histogram representation limits the use of these methods in real-time applications. Other promising work includes the Sobel-LBP approach [43] and the Local Derivative Pattern (LDP) [39].

Recently researchers presented LBP-like descriptors based on discriminative learning. Notable examples include Decision Tree-based LBP (DTLBP) [25], Discriminative LBP (DLBP) [26], Local Quantized Patterns (LQP) [14] and Discriminative Face Descriptor (DFD) [17]. DTLBP attempts to learn discriminative patterns using decision tree induction, but one drawback of the DTLBP is the high cost of constructing and storing the decision trees, especially when large neighborhoods are used. The DLBP method represents LBP-like descriptors as a set of pixel comparisons between a center pixel and its 48 neighboring pixels in a 7×7 patch, and seeks a subset from among all comparisons that maximizes the Fisher discrimination power. Unfortunately, this is an intractable combinatorial optimization problem. Conversely, LQP makes use of vector quantization and lookup tables to allow local pattern features a larger neighborhood and more quantization levels. LQP adopts the neighborhood of Fig. 2(a), and attempts to learn a codebook from all the LBP codes that have been extracted. This codebook learning is not feasible with a large neighborhood since the number of local binary patterns grows exponentially with the number of neighboring pixels. A further constraint is the size of the lookup table, which must store the output code of each possible input pattern. Finally, the recent DFD descriptor learns the most discriminating local features, by computing the difference vectors between the center patch and each of its neighboring patches to form a Pixel Difference Matrix, which is then projected and re-grouped to form the discriminant pattern vector. A standard bag-of-words model [21] is applied for face representation. The DFD descriptor performs well in face identification and verification, but involves a number of parameters, has high dimensionality, requires time-consuming training, and has high training data needs.

Enhancing noise robustness

Tan and Triggs [35] introduced Local Ternary Patterns (LTP), where the binary LBP code was replaced by a ternary one. The LTP method was more resistant to noise, but not strictly invariant to gray-scale changes, and the selection of additional threshold values was not easy. Ahonen et al. [2] introduced Soft LBP (SLBP) histograms, which enhanced robustness by incorporating fuzzy membership in the representation of local texture primitives. Fuzzification allows multiple local binary patterns to be generated, and thus one pixel position can contribute to more than a single bin in the histogram of possible LBP patterns. However, this sacrifices invariance to monotonic variations, and computation of the contribution weights creates a computational burden. Noting this disadvantage of SLBP, Ren et al. [32] proposed a much more efficient variant: the Noise Resistant LBP (NRLBP), which has fast contribution weight computation.

Combining approaches

A current trend in the development of new effective local descriptors is combining the strengths of multiple complementary descriptors. Huang et al. [13] proposed the Extended LBP (ELBP) method, which not only performed binary comparison between the central pixel and its neighbors, but also encoded their gray value differences. Motivated by ELBP [9], Guo et al. developed a Completed LBP (CLBP) method by decomposing the local differences into two complementary components, the signs and the magnitudes, which were coded using two LBP-like descriptors. Chan et al. [6] proposed to combine multiscale LBP with multi-scale Local Phase Quantization (LPQ) using kernel fusion for face recognition, and claimed improved performance. Klare et al. [15] studied the combination of LBP and SIFT for heterogeneous face recognition.

3. Extended LBP descriptors

Guo et al. [9] proposed a complete LBP for texture classification, which included both the sign and the magnitude components between a given central pixel and its neighbors in order to improve the discriminative power of the original LBP operator. The operator derived from the sign component, denoted as LBP_S, is the same as the original LBP operator defined in (2):

$$LBP_S_{r,n} = LBP_{r,n}$$



Fig. 3. The two given patterns, left, would be considered equivalent by LBP, although the patterns are, in some ways, quite different from one other, not just due to a rotation, but due to underlying textural properties, revealed via angular and radial differences.

The operator computed from the magnitude component, denoted as LBP_M, performed a binary comparison between the absolute value of the difference between the central pixel and its neighbors and a global threshold to generate an LBP-like code:

$$LBP_M_{r,p} = \sum_{n=0}^{p-1} s(|x_{r,p,n} - x_{0,0}| - \mu_{r,p})2^n$$
(4)

where

$$\mu_{r,p} = \frac{1}{p} \sum_{n=0}^{p-1} |x_{r,p,n} - x_{0,0}|.$$
(5)

3.1. Angular-differences based descriptors

LBP encodes only the relationship between a central point and its neighbors. Moreover, its capability of encoding image configuration and pixel-wise relationships is further weakened by the binary quantization of pixel differences. Although extended to facilitate analysis of image textures at multiple scales [30] by varying the sampling radius (r) and the number of sampling points (p), important information regarding the relationship between neighboring points on the same radius (intra-radius) and the relationship between neighboring points across different radii (inter-radius) was lost.

A simple example is shown in Fig. 3. For standard LBP the two patterns were classified as the same class, although the textural structures represented were actually quite different from one another. More specifically, the top and bottom patterns in Fig. 3 show quite different inter-radius and intra-radius intensity variations, information which cannot be revealed by LBP_S since it encodes only the relationship between a center pixel and its neighbors. With the aim to address these issues, we investigated the feasibility and effectiveness of encoding angular and radial difference information for face representation, where the angular and radial directions were on a circular grid. More specifically, we proposed two families of related descriptors, the Angular Difference-based Local Binary Pattern (ADLBP) and the Radial Difference-based Local Binary Pattern (RDLBP). Similar to LBP_S and LBP_M, both sign and magnitude components of angular and radial differences were considered, leading to four LBP-like descriptors: ADLBP_S, ADLBP_M, RDLBP_S, and RDLBP_M, all illustrated in Fig. 4. The proposed new descriptors were not designed to compete with traditional LBP, but rather complement it and extend the set of feature candidates. The derivation of ADLBP_S and ADLBP_M, were introduced first, followed by RDLBP_S and RDLBP_M in Section 3.2.

For each pixel in an image, we considered the accumulated angular differences computed from its uniformly distributed neighbors in a ring of radius *r*. Similar to the sampling scheme in the original LBP approach, we sampled pixels around a central pixel $x_{0,0}$, but restricted the number of points on any circle of radius *r* sampled to be a multiple of eight, and thus p = 8q for some positive integer *q*. Therefore, the neighbors of $x_{0,0}$ sampled on radius *r* are $\mathbf{x}_{r,8q} = [x_{r,8q,0}, \dots, x_{r,8q,8q-1}]^T$.

Empirical tests suggested that averaging over two adjacent radii produced the best results, so the ADLBP_S descriptor was computed by accumulating the angular differences of radii r and r - 1:

$$\underline{\psi}_{r,\delta,q} = \underline{\Delta}_{r,\delta,q}^{\mathrm{Ang}} + \underline{\Delta}_{r-1,\delta,q}^{\mathrm{Ang}} \tag{6}$$

where

$$\underline{\Delta}_{r,\delta,q}^{\mathrm{Ang}} = \left[\Delta_{r,\delta,q,0}^{\mathrm{Ang}}, \dots, \Delta_{r,\delta,q,8q-1}^{\mathrm{Ang}}\right]^{T}$$
(7)



Fig. 4. The proposed extended set of features, made up of (a) local intensity features, (b) local angular-difference based features, and (c) local radialdifference based features.

$$\Delta_{r,\delta,q,n}^{\text{Ang}} = \mathbf{X}_{r,p,n} - \mathbf{X}_{r,p,\text{mod}(n+\delta,p)}$$
(8)

The latter term is the angular difference computed over a given angular displacement $\delta(2\pi/p)$, where $x_{r, p, n}$ and $x_{r, p, \text{mod}(n+\delta, p)}$ correspond to the values of pixels at radius *r* spaced δ elements apart.

Considering the rapid increase of the number of binary patterns with the increase of sampling points p, we did not want to derive local binary patterns directly from $\underline{\psi}_{r,\delta,q}^{\text{Ang}}$, and instead followed the idea in our recent work in [23]. We transformed $\underline{\psi}_{r,\delta,q}^{\text{Ang}}$ by local averaging along an arc,

$$\phi_{r,\delta,q,n}^{\text{AngSign}} = \frac{1}{q} \sum_{k=0}^{q-1} \psi_{r,\delta,q,(qn+k)}^{\text{Ang}}, \quad n = 0, \dots, 7,$$
(9)

as illustrated in Fig. 4, such that the number of elements in $\phi_{r,\delta,q}^{AngSign}$ is always eight.

Finally, given the sign component vector $\phi_{r,\delta,q}^{\text{AngSign}}$, we trivially computed a binary pattern as in LBP_S in (2):

$$ADLBP_S_{r,\delta,q} = \sum_{n=0}^{7} s(\phi_{r,\delta,q,n}^{AngSign}) 2^n$$
(10)

Similar to LBP_M, we also proposed the ADLBP_M descriptor based on the magnitude component of angular differences $|\underline{\psi}_{r,\delta,q}^{\text{Ang}}|$. Specifically, we transformed the absolute angular differences $|\underline{\psi}_{r,\delta,q}^{\text{Ang}}|$, via local averaging as in (9):

$$\phi_{r,\delta,q,n}^{\text{AngMag}} = \frac{1}{q} \sum_{k=0}^{q-1} |\psi_{r,\delta,q,(qn+k)}^{\text{Ang}}|, \quad n = 0, \dots, 7.$$
(11)

from which we derived ADLBP_M as in LBP_M:

$$ADLBP_M_{r,\delta,q} = \sum_{n=0}^{\prime} s(\phi_{r,\delta,q,n}^{AngMag} - \mu_{r,\delta,q}^{AngMag})2^n,$$
(12)

where $\mu_{r,\delta,q}^{\mathrm{AngMag}}$ is the average of $\{\phi_{r,\delta,q,n}^{\mathrm{AngMag}}\}_{n=0}^7$

3.2. Radial-differences based descriptors

In order to obtain the radial-differences based descriptor RDLBP_S and RDLBP_M, we first computed the radial difference vector:

$$\underline{\Delta}_{r,\delta,q}^{\text{Rad}} = [\Delta_{r,\delta,q,0}^{\text{Rad}}, \dots, \Delta_{r,\delta,q,8q-1}^{\text{Rad}}]^T,$$
(13)

where $\Delta_{r,\delta,q,n}^{\text{Rad}} = x_{r,p,n} - x_{r-\delta,p,n}$ is the radial difference computed with given integer radial displacement δ , and $x_{r,p,n}$ and $x_{r-\delta,p,n}$ correspond to the intensity values of pairs of pixels spaced δ apart in radius but in the same direction, as shown in Fig. 4.

We computed the transformed sign and magnitude components $\underline{\phi}_{r,\delta,q}^{\text{RadSign}}$ and $\underline{\phi}_{r,\delta,q}^{\text{RadMag}}$ by averaging, as before:

$$\phi_{r,\delta,q,n}^{\text{RadSign}} = \frac{1}{q} \sum_{k=0}^{q-1} \Delta_{r,\delta,q,(qn+k)}^{\text{Rad}}, \quad n = 0, \dots, 7$$
(14)

$$\phi_{r,\delta,q,n}^{\text{RadMag}} = \frac{1}{q} \sum_{k=0}^{q-1} |\Delta_{r,\delta,q,(qn+k)}^{\text{Rad}}|, \quad n = 0, \dots, 7.$$
(15)

Similar to (10) and (12), the RDLBP_S and RDLBP_M descriptors are defined as follows:

$$\text{RDLBP}_{S_{r,\delta,q}} = \sum_{n=0}^{7} s(\phi_{r,\delta,q,n}^{\text{RadSign}}) 2^n \tag{16}$$

$$\text{RDLBP}_{M_{r,\delta,q}} = \sum_{n=0}^{7} s(\phi_{r,\delta,q,n}^{\text{RadMag}} - \mu_{r,\delta,q}^{\text{RadMag}})2^n, \tag{17}$$

where $\mu_{r,\delta,q}^{\text{RadMag}}$ is the average of $\{\phi_{r,\delta,q,n}^{\text{RadMag}}\}_{n=0}^7$

3.3. Statistical distribution examination of patterns

In standard face recognition tasks using LBP, only uniform patterns were used [1,5,37]. The nonuniform patterns were considered in only a single bin of the histogram that was used to extract features in the classification stage. However, the representation of texture information using only uniform patterns may be problematic [10,19], since the uniform patterns may not be the dominant patterns, especially when the radius of the LBP operator increased. Fig. 5 shows distributions of the patterns extracted from the gallery faces in the FERET database [31] for the proposed descriptors. The first nine bins of each histogram corresponds to the rotation invariant uniform patterns, and the remaining bins to the nonuniform patterns. It is clear that the uniform patterns may not necessarily represent a significant proportion of overall patterns. This was most striking in the cases of ADLBP_S and ADLBP_M. Consequently, textural information could not be effectively represented by solely considering the histogram of the uniform patterns.

The original uniform patterns introduced by Ojala et al. [30] did not represent most of the observed patterns, and thus we proposed use of the full descriptors, *i.e.* 256 binary patterns for each descriptor at each scale. Furthermore, as pointed out by Ojala et al. [30], the occurrence frequencies of different patterns vary greatly and some of the patterns rarely occur, as illustrated in Fig. 5. Therefore, we proposed a labeled dominant pattern (LDP) scheme which learned the most frequently occurring patterns from a training set to reduce the dimensionality of the proposed descriptors without information loss. First, the pattern histogram vectors of all images in the training set were computed. Then they were combined into one histogram and sorted in descending order, and the labels of the minimum set of patterns that occupied 80% of the total pattern occurrences were retained. At last, given the learned set of dominant patterns, any image was represented by the probabilities of the dominant patterns. In summary, the pseudo-codes on determining the LDP patterns of an LBP descriptor is presented in Fig. 6.

The proposed LDP scheme is different from the DLBP approach proposed by Liao et al. [19]. In DLBP, for each image $I_i \in S^{\text{train}}$ in the training set, the pattern histogram was extracted and sorted in descending order. Then the smallest number of patterns k_i for 80% pattern occurrences was found. Finally, the average number $k = \frac{1}{|S^{\text{train}}|} \sum k_i$ was defined as the required number of patterns that occupy dominant ratio. The DLBP approach was not limited to consider only a fixed set of patterns. For two different texture images, the dominant patterns could be of different types. It is important to remark that this is an unlabeled feature reduction scheme in which no information is retained about the patterns' labels, and some important information was lost. In contrast, in our approach, the strong pattern occurrences were defined *jointly* for the entire image set, such that a consistent set of features was extracted for each image, and thus each feature dimension corresponded to a fixed pattern type.













Fig. 5. Proportions of the rotation invariant LBPs for the proposed LBP-like descriptors: LBP_S_{4, 8}, LBP_M_{4, 8}, ADLBP_S_{4, 1, 4}, ADLBP_M_{4, 1, 4}, RDLBP_S_{4, 1, 4} and RDLBP_M_{4, 1, 4} using the FERET gallery set.

Input: The gallery face set, and the threshold parameter
$0 < \tau \leq 1$ to determine a set of labeled dominant
patterns (LDP) for an LBP type descriptor.
Output : LDP pattern set S_{LDP}
Initialization : Initialize the original pattern histogram \underline{h}
where $h_i = 0$ and the pattern index vector
\underline{v} , where $v_i = i - 1, i = 1, \cdots, d$ where d
denotes the number of all possible patterns
types of interest (e.g. $d = 2^8 = 256$).
for each face image I in the gallery set do
for each pixel x_k from image I do
Compute the binary pattern label of x_k , $l(x_k)$
Increase the corresponding bin by 1, $h_{l(x_k)}$ + +
end
end
Sort \underline{h} in descending order, resulting $\underline{h}^{\text{sort}}$. Change the
configuration of the pattern index vector \underline{v} according to
the element order of $\underline{h}^{\text{sort}}$, leading to vector $\underline{v}^{\text{sort}}$.
for $k \leftarrow 1$ to d do
if $(\sum_{i=1}^{k} h_i^{\text{sort}}) / (\sum_{i=1}^{d} h_i^{\text{sort}}) > \tau$
break
end
end
Return $\mathcal{S}_{\text{LDP}} = \{v_1^{\text{son}}, \cdots, v_k^{\text{son}}\}.$

L. Liu et al./Information Sciences 358-359 (2016) 56-72



Fig. 7. Overview of the proposed face recognition framework.

4. Face representation

With the proposed descriptors detailed in Section 3, our face feature representation pipeline is similar to the one proposed by Ahonen et al. [1], but incorporates a more sophisticated illumination normalization step as used by Tan and Triggs [35]. Fig. 4 illustrates the proposed extended set of features, and Fig. 7 summarizes the main operation for the face recognition pipeline. Histogram feature extraction from a face image involves the following steps:

- (1) Crop the face region and align the face by mapping the eyes to a canonical location with a similarity transform.
- (2) Photometrically normalize the cropped faces by utilizing the preprocessing sequence approach [35].
- (3) Divide the face image into a non-overlapping grid with equally sized subregions.
- (4) For each subregion, apply a feature extraction operator (*e.g.* ADLBP_S) to each pixel. Afterward, create a histogram of the feature values and concatenate these histograms into a single feature vector to represent the face image.

4.1. Histogram features

In order to incorporate more spatial information into the final descriptor, the proposed feature images were spatially divided into multiple non-overlapping regions and histograms were extracted from each region. More precisely, a face image I_i is divided into $\sqrt{w} \times \sqrt{w}$ non-overlapping sub-regions, from each of which a sub-histogram feature of dimensionality m is extracted and is normalized to sum one. By concatenating these regional sub-histograms into a single vector, a final histogram feature \underline{h}_i^o of dimensionality n = mw is obtained for face representation. We propose to compare three choices for the extracted features: The uniform ("u2") patterns, all of the patterns ("full"), and the LDP scheme ("ldp") as just discussed in Section 3.3.

Let \underline{h}_{1i}^{o} , \underline{h}_{2i}^{o} , \underline{h}_{3i}^{o} , \underline{h}_{4i}^{o} , \underline{h}_{5i}^{o} and \underline{h}_{6i}^{o} denote the histogram feature produced by descriptor LBP_S, LBP_M, ADLBP_S, ADLBP_M, RDLBP_S and RDLBP_M respectively, as illustrated in Fig. 7.

4.2. Dimensionality reduction and feature fusion

The proposed extended set of LBP descriptors provide rich information, but need to be used to construct a powerful classifier. Recent methods frequently combine multiple local features such as LBP, LTP, SIFT, LPQ and Gabor [6,15,35], where the actual combination strategy ranges from simple summing at the decision level to multiple kernel learning. For the specific problem of face recognition, linear projection models [28,37] and kernel LDA [35] have been used. We strived to show that the proposed radial and angular difference descriptors, which are designed to encode additional types of local texture information, do capture information complementary to that of intensity-based descriptors and thus, when combined, improve overall face recognition performance.

A straightforward way to fuse the proposed features is to concatenate the feature vectors. Since the dimensionality n of a single histogram feature $\underline{\mathbf{h}}_{ki}^{o}(k = 1, ..., 6)$ is already relatively high, a drawback of feature concatenation is the so-called "curse of dimensionality", creating significant burdens on memory, storage and computation. Like other researchers [14,17,28,37], we believe that strong correlations exist between the proposed features and that most of the dimensions carry redundant information, and hence further dimension reduction is necessary.

The standard PCA technique selects a dimensionality reducing linear projection that maximizes the scatter of all projected samples, which means the scatter being maximized is due not only to the interclass scatter that is useful for classification, but also to the intraclass scatter unwanted for classification purposes. PCA has two disadvantages: (1) the leading eigenvectors primarily encode variations such as illumination and expression, rather than information that is useful for discrimination [3]; and (2) the Mean-Square-Error principle underlying PCA favors low frequencies, and results in the loss of discriminative information contained in high frequency components.

The idea behind WPCA [14,17,28,37] is that discriminant information is equally distributed along all the principle components. Therefore, all the principal components are divided by the square-roots of the corresponding eigenvalues to normalize the projected features to the same variance. Specifically, suppose we have the PCA projected feature $\underline{y} = W_{pca}^T \underline{h}^0$, where W_{pca} is the projection matrix computed by the standard PCA technique and is a matrix with orthonormal columns. Then, \underline{y} is subjected to a whitening transformation and yields the final WPCA projected feature \underline{h} :

$$\underline{h} = \mathbf{\Lambda}^{-1/2} \underline{y} = \mathbf{\Lambda}^{-1/2} \mathbf{W}_{\text{pca}}^T \underline{h}^o$$
(18)

where $\Lambda^{-1/2} = diag\{\lambda_1^{-1/2}, \lambda_2^{-1/2}, \ldots\}$, with λ_i being its eigenvalue. The integrated projection matrix $\Lambda^{-1/2} \mathbf{W}_{pca}^T$ treats variance along all principal component axes as equally significant by more heavily weighting components corresponding to smaller eigenvalues and is arguably appropriate for discrimination. Consequently, the negative influences of the leading eigenvectors are reduced, whereas the discriminating details encoded in trailing eigenvectors are enhanced [14].

We applied WPCA to each single feature, letting $\underline{\mathbf{h}}_{ki}$ be the WPCA projected feature of $\underline{\mathbf{h}}_{ki}^{o}$. The similarity function is denoted as $d(\underline{\mathbf{h}}_{ki}, \underline{\mathbf{h}}_{ki})$; for fusing K features, we used the sum of the similarity measures of the individual features:

$$d(\mathbf{I}_i, \mathbf{I}_j) = \sum_{k}^{K} d(\underline{\mathbf{h}}_{ki}, \underline{\mathbf{h}}_{kj}).$$
(19)

As adopted in previous research [14,17,37], WPCA was conducted on the gallery set only.

4.3. Distance measure and classification

To perform face recognition, there are two crucial components: (1) extracting low-level features, and (2) building classification models. In this work, focus was on evaluating the discrimination properties of the proposed descriptors, and therefore tried to make as few assumptions as possible and chose the simple Nearest Neighbor Classifier (NNC).

First, we evaluated the effectiveness of the original histogram feature vector $\{\underline{h}_{ki}^{o}\}_{k=1}^{\delta}$ corresponding to the proposed descriptors. The distance between two normalized frequency histograms was measured using the Chi squared distance. The face samples were then classified according to their normalized histogram feature vectors \underline{h}_{i}^{o} and \underline{h}_{i}^{o} :

$$\chi^{2}(\underline{\mathbf{h}}_{i}^{o},\underline{\mathbf{h}}_{j}^{o}) = \frac{1}{2} \sum_{k} \frac{[\underline{\mathbf{h}}_{i}^{o}(k) - \underline{\mathbf{h}}_{j}^{o}(k)]^{2}}{\underline{\mathbf{h}}_{i}^{o}(k) + \underline{\mathbf{h}}_{j}^{o}(k)}$$
(20)

the same distance metric used in [1].

We further evaluated the performance of the more efficient features $\{\underline{\mathbf{h}}_{ki}\}_{k=1}^{6}$, which were obtained by applying the WPCA technique on the original histogram feature vectors $\{\underline{\mathbf{h}}_{ki}^{0}\}_{k=1}^{6}$. The WPCA projected features were no longer histogram frequencies, and therefore the χ^2 distance may not be a good choice. Instead, we employed the conventional Euclidean distance metric to compare faces in the normalized projected space.

5. Experiments

We conducted extensive experiments on three commonly-used databases: Extended Yale B, FERET and CAS-PEAL-R1. We used the standard evaluation protocol for each database in order to facilitate comparison with previous work. These databases incorporate several deviations from the ideal conditions, including illumination, expression, occlusion and pose alterations. We adopt several of the standard evaluation protocols reported in the face recognition literature, and we present a comprehensive comparison of the proposed approach with the recent state of the art.

5.1. Image data and experimental setup

The **Extended Yale B** database [16] consists of 2414 frontal face images of 38 subjects, with about 64 samples per subject, which were captured under controlled lighting conditions and were cropped and normalized to a standard size of 192×168 pixels. Instead of randomly selecting half of the images per subject for training and the remaining for testing, we used a more difficult setup [27,34]. The database was divided into five subsets according to light direction with respect to the camera axis (Fig. 8(a)): subset 1 consisted of gallery images under nominal lighting conditions, while all others were used as probes. Subsets 2 and 3 were characterized by slight-to-moderate illumination variations, while subsets 4 and 5 depicted severe illumination changes.

The **FERET** database [4,31] is one of the largest publicly available databases, consisting of a total of 14,051 gray scale images representing 1199 subjects. The images contain variations including lighting, facial expressions and pose angle. To



Fig. 8. (a) Examples of one of the 38 subjects in the Extended Yale database B under the five subsets of lighting [16]. (b) Cropped example face images from the FERET database [4,31]. (c) Cropped example face images from CAS-PEAL-R1 [8].

ensure the reproducibility of tests, we used the publicly available CSU face identification evaluation framework [4] to evaluate the performance of our proposed methods, with a standard gallery (fa) and four probe sets (fb, fc, dup1, dup2). All images were normalized according to the provided eye coordinates and cropped to 150×130 . Example images are shown in Fig. 8(b).

The **CAS-PEAL-R1** database [8] is a large-scale Chinese face database for face recognition algorithm training and evaluation, containing 30,863 images of 1040 individuals. The standard experimental protocol [8] divides the data into a training set, a gallery set and six probe sets. No overlap exists between the gallery and any of the probes. The gallery contains one image taken under standard conditions for each of the 1040 subjects, while the six probes respectively contain images with the following variations: expression, lighting, accessories, background, distance and aging. Like in previous work [17], we used the expression, lighting, accessory subsets as probes. All images were cropped to 150×130 based on the provided eye coordinates, with examples shown in Fig. 8(c). The CAS-PEAL-R1 is challenging because it contains 27 times more subjects than in the Extended Yale B, and has a greater degree of intrinsic variability due to its more natural image capture conditions.

For all databases, all face samples were photometrically normalized by the preprocessing sequence approach proposed by Tan and Triggs [35]. This photometric normalization method reduced the effects of illumination variation, local shadowing and highlights, while still keeping the essential visual appearance information for use in recognition.

5.2. Parameter evaluation

The key parameters involved in the proposed method are as follows:

- 1. number of sampling points 8q on radius r;
- 2. the step parameter δ for computing differences;
- 3. uniform vs. full vs. LDP;
- 4. Chi square+NNC vs. WPCA+Euclidean+NNC;
- 5. which descriptors to fuse;
- 6. the number of sub-regions $\sqrt{w} \times \sqrt{w}$;
- 7. the radius *r* for extracting the proposed descriptors.

For the LBP_S and LBP_M, we used a constant q = 1 (*i.e.* a constant number of 8 sampling points for any radius r), as done in previous work [1,5,6]. For other descriptors, we proposed an intermediate sampling scheme aimed to prevent over-smoothing at larger radii r, and set q = r for $r \le 4$, and q = 4 for r > 4. In terms of parameter δ , we set $\delta = 1$.

Table 1 shows the recognition rates of the individual proposed descriptors on the Extended Yale B database, comparing the three pattern grouping schemes discussed in Section 3.3, using only the elementary histogram features (without dimensionality reduction by WPCA).

From Table 1, we can make the following observations. First, for each proposed descriptor, the histogram representation generated by the full pattern method is ideal, significantly outperforming the uniform pattern method and marginally outperforming the dominant scheme, clearly demonstrating the insufficiency of the uniform patterns for representing face images. On this basis we will not report results for the uniform pattern scheme in our latter experiments. Second, the

Table 1

Recognition rates (%) of the proposed single elementary features on the Extended Yale B database, comparing "uniform", "full" and "labeled dominant". The NNC classifier with χ^2 distance measure is used for classification. The radius *r* is 4 and the number of subregions $\sqrt{w} \times \sqrt{w}$ is set as 9 × 9.

	Method	S2	S3	S4	S5	Mean
U_1 U_2 U_3 U_4 U_5 U_6	LBP_S ^{u2} LBP_M ^{u2} ADLBP_S ^{u2} ADLBP_M ^{u2} RDLBP_S ^{u2} RDLBP_M ^{u2}	99.8 99.8 99.8 99.8 99.8 99.8	99.6 99.2 89.5 99.6 99.4 99.6	93.2 95.8 28.5 94.5 91.9 98.2	77.7 91.7 12.5 88.7 68.5 91.5	92.6 96.6 57.6 95.7 89.9 97.3
F ₁ F ₂ F ₃ F ₄ F ₅ F ₆	LBP_S ^{full} LBP_M ^{full} ADLBP_S ^{full} ADLBP_M ^{full} RDLBP_S ^{full} RDLBP_M ^{full}	99.8 99.8 99.8 99.8 99.8 99.8 99.8	99.8 99.6 99.6 99.6 99.6 99.6	99.6 99.6 91.4 99.4 98.7 98.9	96.2 97.6 67.1 97.8 86.6 95.7	98.9 99.2 89.5 99.2 96.2 98.5
$D_1 \\ D_2 \\ D_3 \\ D_4 \\ D_5 \\ D_6$	LBP_S ^{ldp} LBP_M ^{ldp} ADLBP_S ^{ldp} ADLBP_M ^{ldp} RDLBP_S ^{ldp} RDLBP_M ^{ldp}	99.8 99.8 99.8 99.8 99.8 99.8 99.8	99.6 99.8 99.4 99.6 99.6 99.6	98.3 99.1 90.1 98.9 97.1 98.7	92.7 97.3 65.1 96.5 82.9 94.8	97.6 99.0 88.6 98.7 94.9 98.2
	PCA [27] LRC [27] LRC_Fused [34]	98.5 100 100	80.0 100 100	15.8 83.27 88.97	24.4 33.61 84.73	54.7 79.2 93.4

Table 2

Recognition rates (%) of the proposed WPCA features on the FERET database. The WPCA dimension is 1196 (the size of the gallery set). The NNC classifier with Euclidean distance measure is used for classification. The values of parameters *r* and *w* are the same as those in Table 1.

	Method	fb	fc	dup1	dup2
F ₁	LBP_S ^{full} + χ^2	96.1	97.4	75.2	70.1
F ₂	$LBP_M^{full} + \chi^2$	92.8	95.9	70.4	66.2
F ₃	ADLBP_S ^{full} + χ^2	97.2	97.4	75.9	73.1
F ₄	ADLBP_M ^{full} + χ^2	95.7	97.9	73.3	70.9
F ₅	$RDLBP_S^{full} + \chi^2$	95.7	97.9	73.3	70.9
F ₆	RDLBP_M ^{full} + χ^2	91.5	96.9	66.2	66.7
F ₁	LBP_S ^{full} +WPCA	99.0	100.0	83.2	80.3
F ₂	LBP_M ^{full} +WPCA	98.3	99.0	79.4	78.6
F ₃	ADLBP_S ^{full} +WPCA	97.1	95.9	78.9	73.9
F ₄	ADLBP_M ^{full} +WPCA	98.1	98.5	75.1	73.9
F ₅	RDLBP_S ^{full} +WPCA	86.9	79.9	50.6	45.7
F ₆	RDLBP_M ^{full} +WPCA	93.1	86.6	58.7	52.1

magnitude-based descriptors consistently outperformed the corresponding sign-based descriptors, particularly under very severe lighting changes (S4 and S5). This finding is different from those by Guo et al. [9], in that the sign component was more informative than magnitude. Finally, the proposed LBP_S, LBP_M, ADLBP_M and RDLBP_M (either with full or dominant patterns) outperformed the results obtained using a Linear Regression Classification (LRC) method [27], particularly under severe lighting variations.

Table 2 illustrates the effect of WPCA dimensionality reduction on the FERET database. The principal component axes were calculated using the 1196 gallery faces, with no additional training set. We observed that the recognition performance of the individual descriptors was consistently improved by WPCA+Euclidean relative to Chi-square, with the exception of the radial descriptors RDLBP_S and RDLBP_M, possibly because the WPCA dimensionality of 1196 is not adequate for these two descriptors to represent face images. A more significant consideration was the computational advantage of WPCA+Euclidean+NNC over Chi square+NNC due to the much lower feature dimensionality of WPCA. In our further experiments we will focus on WPCA+Euclidean+NNC.

We did not expect individual descriptors to capture all of the salient image features in recognizing a face. Indeed, it has been our premise, from the outset, to wish to identify and fuse those descriptors having complementary information. To simplify the notation, we refer to the full and LDP features as F_i and D_i , respectively, as identified in Table 1 and 2.

It would be possible to directly fuse all the six proposed descriptors, but there are two concerns. First, it is not obvious that fusing all six features necessarily gives the best recognition performance, and second, there is a concern regarding processing and computational costs. Other things being equal, we would prefer to fuse the fewest features giving the strongest performance.

Table 3

Recognition rates (%), comparing results of fusing multiple descriptors. WPCA dimension set as 1196 for FERET and 1040 for CAS-PEAL-R1. Values of parameters *r* and *w* are the same as those in Table 1.

Method	FERET					CAS-PEAL-R1			
	fb	fc	dup1	dup2	Mean	Expression	Accessary	Lighting	Mean
F ₁	99.0	100.0	83.2	80.3	92.4	97.5	92.4	42.9	75.5
F ₂	98.3	99.0	79.4	78.6	90.6	94.1	85.1	36.5	69.5
F ₃	97.1	95.9	78.9	73.9	89.1	98.1	93.6	47.0	77.6
F ₄	98.1	98.5	75.1	73.9	88.6	95.1	87.0	42.1	72.6
F ₅	86.9	79.9	50.6	45.7	71.0	96.1	90.5	33.1	70.8
F ₆	93.1	86.6	58.7	52.1	77.9	91.3	78.1	34.8	65.6
F ₃₄	99.4	100.0	85.6	84.2	93.7	98.4	93.5	59.2	82.1
F ₁₄	99.4	100.0	87.3	86.3	94.4	98.3	93.6	57.6	81.6
F ₁₃	99.2	100.0	87.7	84.6	94.3	98.7	94.7	55.8	81.4
F ₁₃₄	99.6	100.0	88.8	87.2	95.1	98.9	94.8	63.4	84.3
F ₁₃₄₆	99.6	100.0	88.9	87.6	95.1	98.4	93.9	65.9	84.8
F ₁₃₄₅₆	99.6	100.0	89.1	86.8	95.1	98.7	94.2	66.4	85.1
F ₁₂₃₄₅₆	99.6	100.0	88.8	87.2	95.1	98.5	93.8	66.2	84.8

Table 4

Comparisons of the recognition rates (%) of fusing the proposed LDP features on the FERET and CAS-PEAL-R1 Databases.

Method	FERET					CAS-PEAL-R1			
	fb	fc	dup1	dup2	Mean	Expression	Accessary	Lighting	Mean
D ₁	98.8	100.0	81.4	77.8	91.5	96.6	90.8	34.9	71.7
D ₂	98.4	97.9	78.1	75.2	89.8	93.0	83.1	33.4	67.3
D ₃	97.2	98.5	80.9	77.8	90.4	97.5	92.1	41.2	74.8
D_4	98.2	98.5	75.6	73.9	88.8	94.3	84.2	37.0	69.4
D ₅	94.4	95.9	63.6	58.5	81.4	94.9	88.4	27.6	67.7
D ₆	96.2	94.3	68.7	65.4	84.5	89.9	75.9	31.2	63.1
D ₃₄	99.3	99.5	86.8	84.6	94.0	98.0	92.8	55.3	80.4
D ₁₄	99.3	100.0	86.4	84.6	93.9	98.0	92.3	51.5	78.8
D ₁₃	98.8	99.5	87.5	83.8	93.9	98.3	93.8	50.8	79.2
D ₁₃₄	99.3	100.0	88.6	86.8	94.8	98.7	94.5	60.3	83.0
D ₁₃₄₆	99.5	100.0	89.1	87.6	95.1	98.2	93.6	63.9	83.8
D ₁₃₄₅₆	99.6	100.0	89.6	87.6	95.4	98.5	94.1	64.3	84.3
D ₁₂₃₄₅₆	99.5	100.0	89.5	86.8	95.2	98.4	93.9	65.4	84.6

We tested all 57 possible descriptor sets, both full and dominant, with experiments carried out on the FERET and CAS-PEAL-R1 databases. For the CAS-PEAL-R1 database, the principal component axes were calculated using the 1040 gallery images, and the WPCA dimension was fixed to 1040. Table 3 compares the results of the feature fusion methods with those achieved by the single descriptor method. Table 3 lists only those feature combinations that performed the best on both the FERET and CAS-PEAL-R1 databases.

The following observations can be made from Table 3. First, the fused features consistently achieved much better recognition scores than any single descriptor, especially on the FERET dup1, dup2 and the CAS-PEAL-R1 lighting probe sets. For example, the fused feature F_{1346} increased the recognition rate by 23%, 19%, 24% and 31% respectively relative to the LBP_S (F₁), ADLBP_S (F₃), ADLBP_M (F₄) and RDLBP_M (F₆) descriptors alone on the CAS-PEAL-R1 lighting probe set, suggesting that the proposed descriptors indeed captured complementary information. Second, it was observed that among the six proposed descriptors, the LBP_S (F₁), ADLBP_S (F₃), ADLBP_M (F₄) played the more important roles in feature fusion, whereas LBP_M (F₂) played the least important. Finally, the fused features F_{1346} , F_{13456} and F_{123456} gave robust results on the test databases, differing only in computational cost.

Paralleling Table 3, Table 4 lists the results of the fused LDP features. The results in Table 4 support the findings given by Table 3.

Figs. 9 and 10 plot the results of the fused full descriptors as a function of the number of sub-regions $\sqrt{w} \times \sqrt{w}$. In general, the recognition accuracy peaked for an intermediate value of *w*, suggesting that a small number of large regions led to a loss of spatial information, whereas a large number of small regions increased the computation time and degraded system accuracy. Conclusively, 9 × 9 was a defensible choice for both the FERET and CAS-PEAL-R1 databases.

Finally, we addressed the question of radius parameter r with regards to recognition performance. Fig. 11 plots the results of the proposed fused descriptors against the radius parameter r, suggesting that r = 4 and r = 5 are ideal choices.

5.3. Comparative evaluation

With parameter settings determined, examined the effectiveness of the proposed fused descriptors against previously established methods: traditional multiscale LBP by Ahonen et al. [1]; Local Gabor Binary Pattern (LGBP) by Nguyen et al.



Fig. 9. Recognition rates (%) of the WPCA features as a function of the number of subregions on the FERET database.



Fig. 10. Recognition rates (%) of the WPCA features as a function of the number of subregions on the CAS-PEAL-R1 database.



Fig. 11. Recognition rates (%) of the WPCA features as a function of the radius parameter r on the CAS-PEAL-R1 database.

[28], which utilizes Gabor filter bank as preprocessing before LBP extraction; Histogram of Gabor Phase Patterns (HGPP) by Zhang et al. [40], which quantizes Gabor filtered images into histograms that encode not only the magnitude, but also the phase information from the image; Multi-scale LBP Histograms (MLBPH) with LDA learning and kernel fusion by Chan et al. [5]; learning based methods Discriminative LBP (DLBP) [26] and Decision Tree-based LBP (DT-LBP) [25]; and very recent approaches: POEM by Vu et al. published in 2012 [37], which computes LBP-like features from SIFT histograms; LQP by Hussain et al. published in 2012 [14], which first learns significant LQP patterns globally with KMeans clustering and then uses them to encode the face images; DFD by Lei et al. published in 2013 [17], which first learns discriminative local features from local pixel differences and then applies the standard Bag-of-Features (BoF) model to extract histogram features for face representation. The results are shown in Tables 5 and 6.

Table 5 illustrates that the proposed method either matched or significantly outperformed all other approaches. To the best of our knowledge, these are the best results reported to date on the FERET database. The DFD method was the only one to perform similarly, however it is quite complex, involving LDA learning and KMeans clustering, both of which are time-consuming and require enough training data to extract discriminative and stable face representation. On the contrary, our proposed method is simple, efficient and free of pre-training. Since the face images in the FERET probe sets contain several sources of variation, such as expression, lighting, and aging, these comparisons illustrated that our proposed methods is impressively robust to these extrinsic imaging conditions.

Table 5

Comparisons of the recognition rates (%) of the proposed methods with previous results on the FERET Database. WPCA dimension is 1196. The values of parameters r and w are identical to those in Table 1.

Method	fb	fc	dup1	dup2
LBP [1]	97.0	79.0	66.0	64.0
HGPP [40]	97.5	99.5	79.5	77.8
DT-LBP [25]	99.0	100.0	84.0	80.0
DLBP [26]	99.0	99.0	86.0	85.0
MLBPH+LDA+SUM [5]	99.2	99.5	88.4	85.5
Gabor+WPCA [7]	96.3	99.5	78.8	77.8
LGBP+WPCA [28]	98.1	98.9	83.3	81.6
POEM+WPCA [37]	99.6	99.5	88.8	85.0
LPQ+WPCA [14]	99.8	94.3	85.5	78.6
DFD+WPCA [17]	99.3	99.0	88.8	87.6
F ₁₃₄₆ (proposed)	99.6	100.0	88.9	87.6
F ₁₃₄₅₆ (proposed)	99.6	100.0	89.1	86.8
F ₁₂₃₄₅₆ (proposed)	99.6	100.0	88.8	87.2
D ₁₃₄₆ (proposed)	99.5	100.0	89.1	87.6
D ₁₃₄₅₆ (proposed)	99.6	100.0	89.6	87.6
D ₁₂₃₄₅₆ (proposed)	99.5	100.0	89.5	86.8

Table 6

Comparisons of the recognition rates (%) of the proposed methods with previous results on the CAS-PEAL-R1 Database. WPCA dimension is 1040.

Method	Expression	Accessary	Lighting
HGPP [40]	96.8	92.5	62.9
DT-LBP [25]	98.0	92.0	41.0
DLBP [26]	99.0	92.0	41.0
DFD+WPCA [17]	99.0	96.9	63.9
F [°] ₁₃₄₆ (proposed)	98.3	94.2	68.7
F_{13456}^{\diamond} (proposed)	98.5	94.4	70.0
F_{123456}° (proposed)	98.3	94.0	69.6
F [*] ₁₃₄₆ (proposed)	98.4	93.9	72.2
F_{13456}^* (proposed)	98.7	94.4	71.9
F_{123456}^* (proposed)	98.5	94.0	72.3
D_{1346}^{*} (proposed)	98.2	93.0	71.3
D_{13456}^{*} (proposed)	98.5	93.8	71.7
D [*] ₁₂₃₄₅₆ (proposed)	98.4	93.7	72.9

Similarly, Table 6 compares the results of the proposed method with previously reported results on the CAS-PEAL-R1 database. Methods F^{\diamond} were obtained by combining the fused descriptors from two scales, radius r = 4 and r = 5. Methods F^* and D^* used only one scale r = 4, and were then obtained by optimizing the number of subregions $\sqrt{w} \times \sqrt{w}$ over the training set from the CAS-PEAL-R1 database.

As seen in Table 6, our proposed method achieved comparable results on the expression probe set, outperformed all but DFD on the accessory probe set, and significantly outperformed on the lighting probe set. In both accessory and lighting, we outperformed HGPP (which is known to be very computationally expensive) and along with the two learning based methods DT-LBP and DLBP. The high recognition rate on the lighting probe set indicates that the proposed approach is much more robust to illumination variations than other methods in the literature.

As a separate test, we conducted experiments with multiscale F_{13456}° , and with 13×13 subregions, and obtained a recognition rate of **74.6**% on the lighting probe set, which, to the best of our knowledge, is currently the best score achieved on this data set.

6. Conclusions

In this paper we proposed a novel extended set of LBP-like descriptors and developed a simplistic framework to fuse the proposed descriptors for face identification. The proposed set of descriptors consists of two intensity based descriptors LBP_S, and LBP_M, and four accumulated local differences based descriptors ADLBP_S, ADLBP_M, RDLBP_S, and RDLBP_M.

All of the proposed descriptors have desirable features of robustness to lighting, pose, and expression variations, computational efficiency when compared to many of the competing descriptors, and an encoding of both microstructures and macrostructures resulting a more complete image representation. Because the descriptors have a form similar to that of LBP, they inherit the merits of LBP codes and are readily fused to represent face images. The extraction of the proposed descriptors requires no training, and therefore, in contrast to popular learning methods, generalizability is avoided. The proposed local descriptors significantly exploit available information and as well as contain complementary information, which is shown by significantly improved performance of fused descriptors over single descriptor. We found that both full pattern features and the proposed LDP features have strong recognition performance, and the LDP features have lower feature dimensionality.

Results show that our proposed approach outperforms all other descriptor-based techniques, and is comparable with other much more complex learning-based descriptors. High performance coupled with low complexity suggests that our algorithm is a strong candidate for face recognition.

Acknowledgments

This work has been supported by the National Natural Science Foundation of China under contract number 61202336, by the Open Project Program of the National Laboratory of Pattern Recognition (201407354), and the National Basic Research Program of China (973 Program) (2013CB329401).

References

- [1] T. Ahonen, A. Hadid, M. Pietikäinen, Face decription with local binary patterns: application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 28 (12) (2006) 2037–2041.
- [2] T. Ahonen, M. Pietikäinen, Soft histograms for local binary patterns, in: Proceedings of the 2007 Finnish Signal Processing Symposium, 2007, pp. 645–649.
- [3] P. Belhumeur, J. Hespanha, D. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711–720.
- [4] J. Beveridge, D. Bolme, B. Draper, M. Teixeira, The CSU face identification evaluation system its purpose, features, and structure, Mach. Vis. Appl. 16 (2) (2005) 128–138.
- [5] C. Chan, Multi-scale Local Binary Pattern Histograms for Face Recognition, University of Surrey, 2008 Ph.D. thesis.
- [6] C. Chan, M. Tahir, J. Kittler, M. Pietikäinen, Multiscale local phase quantization for robust component-based face recognition using kernel fusion of multiple descriptors, IEEE Trans. Pattern Anal. Mach. Intell. 35 (5) (2013) 1164–1177.
- [7] W. Deng, J. Hu, J. Guo, Gabor-eigen-whiten-cosine: a robust scheme for face recognition, in: Analysis and Modelling of Faces and Gestures, in: Lecture Notes in Computer Science, vol.3723, Springer Berlin Heidelberg, 2005, pp. 336–349.
- [8] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, D. Zhao, The CAS-PEAL large-scale Chinese face database and baseline evaluations, IEEE Trans. Syst. Man Cybern. A Syst. Hum. 38 (1) (2008) 149–161.
- [9] Z. Guo, L. Zhang, D. Zhang, A completed modeling of local binary pattern operator for texture classification, IEEE Trans. Image Process. 19 (6) (2010) 1657–1663.
- [10] Y. Guo, G. Zhao, M. Pietikäinen, Discriminative features for texture description, Pattern Recognit. 45 (10) (2012) 3834–3843.
- [11] X. He, S. Yan, Y. Hu, P. Niyogi, H. Zhang, Face recognition using Laplacian faces, IEEE Trans. Pattern Anal. Mach. Intell. 27 (3) (2005) 328-340.
- [12] D. Huang, C. Shan, M. Ardabilian, Y. Wang, L. Chen, Local binary patterns and its application to facial image analysis a survey, IEEE Trans. Syst. Man Cybern. C Appl. Rev. 41 (6) (2011) 765–781.
- [13] D. Huang, Y. Wang, Y. Wang, A robust method for near infrared face recognition based on extended local binary pattern, in: Advances in Visual Computing, in: Lecture Notes in Computer Science, vol. 4842, Springer Berlin Heidelberg, 2007, pp. 437–446.
- [14] S.u. Hussain, T. Napoléon, F. Jurie, Face recognition using local quantized patterns, in: Proceedings of the British Machine Vision Conference, 2012.
- [15] B. Klare, A. Jain, Heterogeneous face recognition using kernel prototype similarities, IEEE Trans. Pattern Anal. Mach. Intell. 35 (6) (2013) 1410–1422.
 [16] K. Lee, J. Ho, D. Kriegman, Acquiring linear subspaces for face recognition under variable lighting, IEEE Trans. Pattern Anal. Mach. Intell. 27 (5) (2005) 684–698
- [17] Z. Lei, M. Pietikäinen, S. Li, Learning discriminant face descriptor, IEEE Trans. Pattern Anal. Mach. Intell. 36 (2) (2014) 289–302.
- [18] S. Li, A. Jain, Handbook of Face Recognition, second ed., Springer-Verlag, New York, 2011.
- [19] S. Liao, M. Law, A. Chung, Dominant local binary patterns for texture classification, IEEE Trans. Image Process. 18 (5) (2009) 1107–1118.
- [20] S. Liao, X. Zhu, Z. Lei, L. Zhang, S. Li, Learning multi-scale block local binary patterns for face recognition, in: Advances in Biometrics, vol.4642, Springer Berlin Heidelberg, 2007, pp. 828–837.
- [21] L. Liu, P. Fieguth, Texture classification from random features, IEEE Trans. Pattern Anal. Mach. Intell. 34 (3) (2012) 574–586.
- [22] L. Liu, P. Fieguth, G. Zhao, M. Pietikäinen, Extended local binary pattern fusion for face recognition, in: Proceedings of the International Conference on Image Processing, 2014a.
- [23] L. Liu, Y. Long, P. Fieguth, S. Lao, G. Zhao, BRINT: binary rotation invariant and noise tolerant texture classification, IEEE Trans. Image Process. 23 (7) (2014b) 3071–3084.
- [24] L. Liu, L. Zhao, Y. Long, G. Kuang, P. Fieguth, Extended local binary patterns for texture classification, Image Vis. Comput. 30 (2012) 86–99.
- [25] D. Maturana, D. Mery, A. Soto, Face recognition with decision tree-based local binary patterns, in: Proceedings of the Asian Conference on Computer Vision (ACCV), 2010, pp. 618–629.
- [26] D. Maturana, D. Mery, A. Soto, Learning discriminative local binary patterns for face recognition, in: Proceedings of the IEEE International Conference on Automatic Face Gesture Recognition and Workshops (FG), 2011, pp. 470–475.
- [27] I. Naseem, R. Togneri, M. Bennamoun, Linear regression for face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 32 (11) (2010) 2106–2112.
- [28] H. Nguyen, L. Bai, L. Shen, Local Gabor binary pattern whitened PCA: a novel approach for face recognition from single image per person, in: Advances in Biometric, Volume 5558 of the series Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2009, pp. 269–278.
- [29] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures with classification based on feature distributions, Pattern Recognit. 29 (1) (1996) 51–59.
 [20] T. Ojala, M. Pietikäinen, T. Macraët, Multianaluting and the distribution of the
- [30] T. Ojala, M. Pietikäinen, T. Maenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Mach. Intell. 24 (7) (2002) 971–987.
- [31] P. Phillips, H. Moon, S. Rizvi, P. Rauss, The FERET evaluation methodology for face recognition algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 22 (10) (2000) 1090–1104.
- [32] J. Ren, X. Jiang, J. Yuan, Noise-resistant local binary pattern with an embedded error-correction mechanism, IEEE Trans. Image Process. 22 (10) (2013) 4049–4060.
- [33] S. Roweis, L. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science 290 (2000) 2323–2326.
- [34] M. Tahir, C. Chan, J. Kittler, A. Bouridane, Face recognition using multiscale local phase quantisation and linear regression classifier, in: Proceedings of the IEEE International Conference on Image Processing (ICIP), 2011, pp. 765–768.
- [35] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE Trans. Image Process. 19 (6) (2010) 1635–1650.
- [36] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cogn. Neurosci. 3 (1) (1991) 71-86.

- [37] N.S. Vu, A. Caplier, Enhanced patterns of oriented edge magnitudes for face recognition and image matching, IEEE Trans. Image Process. 21 (3) (2012) 1352–1365.
- [38] L. Wolf, T. Hassner, Y. Taigman, Effective unconstrained face recognition by combining multiple descriptors and learned background statistics, IEEE Trans. Pattern Anal. Mach. Intell. 33 (10) (2011) 1978–1990.
- [39] B. Zhang, Y. Gao, S. Zhao, J. Liu, Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor, IEEE Trans. Image Process. 19 (2) (2010) 533–544.
- [40] B. Zhang, S. Shan, X. Chen, W. Gao, Histogram of Gabor phase patterns (HGPP): a novel object representation approach for face recognition, IEEE Trans. Image Process. 16 (1) (2007) 57–68.
- [41] W. Zhang, S. Shan, W. Gao, X. Chen, H. Zhang, Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), vol. 1, 2005, pp. 786–791.
- [42] W. Zhao, R. Chellappa, P. Phillips, A. Rosenfeld, Face recognition: A literature survey, ACM Comput. Surv. 35 (4) (2003) 399–458.
- [43] S. Zhao, Y. Gao, B. Zhang, Sobel-LBP, in: Proceedings of the IEEE International Conference on Image Processing (ICIP), 2008, pp. 2144–2147.