Face recognition under illumination variations based on eight local directional patterns

Mohammad Reza Faraji, Xiaojun Qi

Department of Computer Science, Utah State University, Logan, UT 84322-4205, USA
E-mail: Mohammadreza.Faraji@aggiemail.usu.edu

Abstract: Face recognition under varying illumination is a challenging task. This study proposes a modified version of local directional patterns (LDP), eight LDP (ELDP), to produce an illumination insensitive representation of an input face image. The proposed ELDP code scheme uses Kirsch compass masks to compute the edge responses of a pixel’s neighbourhood. Then, ELDP uses all the directional numbers to produce an illumination invariant image. The author’s extensive experiments show that the ELDP technique achieves an average recognition accuracy of 98.29% on the CMU-PIE face database and 100% on the Yale B face database and clearly outperforms the state-of-the-art representative techniques.

1 Introduction

Face recognition with its wide applications in security, forensic investigation and law enforcement has attracted much attention in the past decade [1]. Illumination variations, pose changes, facial expressions, age variations and occlusion are key factors that make practical face recognition a challenging task [2, 3]. Among these factors, varying lighting conditions in face images are crucial problems to deal with [1]. In the last decade, a large number of illumination preprocessing methods have been proposed. They could be grouped into three categories: grey-level transformation, gradient or edge extraction and face reflection field estimation [1].

Grey-level transformation methods use linear or non-linear functions to redistribute the intensities and correct the uneven illumination to some extent. Representative methods include histogram equalisation and Gamma Intensity Correction [4, 5]. Gradient or edge extraction methods compute gradients of a face image as an illumination-insensitive representation since gradients are more stable than the pixel intensities under different lighting conditions [1]. Some examples of such methods include directional grey-scale derivative (DGD) [6], local binary patterns (LBP) [7] and its modified versions local ternary patterns (LTP) [8] and local Gabor binary pattern histogram sequence [9], local directional patterns (LDP) [10], enhanced LDP (EnLDP) [11], local directional number (LDN) patterns [12], discriminant face descriptor (DFD) [13] and logarithmic fractal dimension (LFD) [14]. DGD extracts grey-level gradients in horizontal and vertical directions which are robust to absolute intensity variations but are still greatly affected by severe shadows. LBP and LTP take a local neighbourhood around each pixel and threshold the neighbourhood pixels based on the value of the central pixel. LDP, EnLDP and LDN use Kirsch compass masks to produce eight directional edge images which are used to produce illumination invariant representations. DFD is an improvement of LBP like feature descriptor. It is involved with three-step feature extraction to maximise the appearance difference from different persons and minimise the difference from the same person. LFD performs a log transformation and then transfers the image to the fractal dimension domain using the fractal analysis to produce the illumination invariant face representations.

Methods in the third category estimate the reflectance field from an image that is illumination invariant. They usually use a face imaging model such as the Lambertian reflectance model in which each pixel of an image is represented as the product of its reflectance and illuminance components [1, 15]. The examples in this category include logarithmic total variation [16], adaptive smoothing (AdaS) [17], gradientface [18] and Weberface [15] methods. For instance, the AdaS method estimates the illumination by smoothing the input image using an iterative method. The gradientface method computes the ratio of gradient 

$\frac{\partial g}{\partial x}$ to the gradient of an image $I(x, y)$ to achieve illumination insensitivity. The Weberface method, inspired by Weber’s law and based on Weber local descriptor [19], produces an illumination insensitive representation by computing a ratio of the local intensity variation to the background.

This paper proposes an illumination invariant face recognition method, called eight local directional patterns (ELDP). This new ELDP method uses all the eight directional numbers instead of the three prominent directional numbers as used in LDP or the top two positive directional numbers as used in EnLDP or the top two positive and negative directional numbers as used in LDN to encode more local structural information and intensity changes of a
face image. It achieves the best face recognition accuracy when compared to seven state-of-the-art methods on two publicly available databases. The proposed ELDP makes the following contributions:

1. Operating in the gradient domain to be robust against illumination variations and noise.
2. Extracting more edge information by utilising all eight directional edge images.
3. Considering relations among eight directional numbers to use valuable structural information.

The rest of this paper is organised as follows: Section 2 proposes the ELDP coding scheme. Section 3 evaluates the performance of the proposed method and shows the experimental results compared with seven state-of-the-art methods. It also presents a method to compute an evaluation criterion, which offers an across comparison of average between-subjects and within-subjects, for each method. This evaluation criterion indicates that the proposed ELDP method should achieve the best face recognition accuracy compared with seven representative methods. Finally, conclusions are given in Section 4.

2 Methodology

2.1 Previous work

Recently proposed LBP, LTP, LDP, EnLDP and LDN approaches produce illumination invariant features. LBP and LTP take a local neighbourhood (e.g. P pixels in a circle of radius of R where P and R are usually set to be 8 and 1, respectively) around each pixel and threshold the neighbourhood pixels based on the value of the central pixel. The LBP approach converts the neighbourhood pixels to the values of 1’s and 0’s and then transfers the 8-bit binary code (or pattern) to the corresponding decimal value. The LTP approach converts the neighbourhood pixels to the values of −1’s, 0’s and 1’s and then splits the ternary pattern into two binary patterns. The final results are LBP and LTP images, which are illumination invariant.

On the other hand, LDP, EnLDP and LDN use Kirsch compass masks ($M_0$, $M_1$, ..., $M_7$) to produce eight directional edge images for a face image, since edge magnitudes are insensitive to lighting variations [12]. One Kirsch mask ($M_0$) is rotated 45° apart in eight directions to produce eight masks as shown in Fig. 1 [20]. The 4th edge direction is computed by the convolution of the original face image $I(x, y)$ and the mask $M_0$. Then, for each pixel, LDP considers an 8-bit binary code so that three bits corresponding to the three prominent numbers out of the eight directional numbers are 1’s and the other five bits are 0’s. LDP considers a 6-bit binary code so that the first three bits code the position of the top positive directional number and the next three bits code the position of the top negative directional number. The generated codes are then converted to their corresponding decimal values to produce LDP, EnLDP and LDN images.

2.2 Eight LDE

To produce the illumination invariant representation, we first use Kirsch compass masks to compute the edge responses (i.e. eight directional edge images). Each of these eight edge images represents the edge significance in its respective direction [12]. These edge responses have different importance levels, but all of them are significant. As a result, we use all the eight directional numbers of each pixel to assign an associated 8-bit binary code. If the directional number of its edge image is positive, we set the value of the respective bit to 1. Otherwise, we set the value of the respective bit to 0. Finally, we compute the corresponding decimal value of the binary code and consider it as the pixel’s ELDP value. Fig. 2 illustrates the detailed steps to compute the ELDP value at one location. The same procedure can be applied to all locations in a face image to produce the final ELDP image as shown in Fig. 2d.

ELDP is similar to LBP in one perspective since both transform the 8-bit binary code to a decimal value. However, they use different mechanisms to generate the 8-bit code. Specifically, ELDP uses the positive or negative edge response information of the eight directional edge images to generate the 8-bit code. LBP uses the sparse points compared with the centre pixel in each neighbourhood of the original image as the threshold (i.e. few number of intensities in a neighbourhood) to generate the 8-bit code. As a result, LBP discards most of the information in the neighbourhood, makes the method sensitive to noise, and limits the accuracy of the method [12]. ELDP avoids the above shortcomings of LBP by using more edge information from the entire neighbourhood and is therefore superior to LBP.

The proposed ELDP is also similar to LDP, EnLDP and LDN methods [10, 12] from the perspective of the use of Kirsch compass masks to compute the edge responses. However, LDP, EnLDP and LDN, respectively, use the three prominent directional numbers, the first and second top positive directional numbers, and the top positive and negative directional numbers to generate their corresponding binary codes. ELDP uses all eight directional numbers to generate its 8-bit binary code. Since these eight directional numbers provide the ‘panorama view’ of gradient directions in a chosen neighbourhood, ELDP offers the following advantages: (i) it is robust against illumination variations and noise by operating in the gradient domain [12], (ii) it extracts more edge information by utilising all eight directional edge images, (iii) it indicates valuable structural information from the neighbourhood by considering the relations among eight directional edge responses and (iv) it is simple and easy to compute.

2.3 Implementation

Similar to other work, we first smooth the face image with a Gaussian filter to reduce the side-effect of shadow boundaries and make the gradient computation more robust [15, 18]. We then implement the ELDP code scheme. The algorithmic view of the proposed method is summarised in Algorithm 1 (see Fig. 3).
Fig. 4 shows sample original images from the CMU-PIE face database and their corresponding ELDP images. It clearly shows the ELDP code scheme reduces the effect of illumination and produces illumination insensitive images since all ELDP images look alike as shown in Fig. 4b.

3 Experimental results

3.1 Experimental settings

The proposed ELDP code scheme is evaluated by conducting experiments on two publicly available CMU-PIE and Yale B face databases with large illumination variations [21, 22]. The Yale B face database has its corresponding publicly available manually cropped database. To be consistent with this database, we also manually crop the PIE face images. We resize all the cropped images to the dimension of 100 × 100 pixels. We compare our method with several recently proposed state-of-the-art methods such as AdaS, Gradientface, Weberface, LBP, LDP, EnLDP and LDN. We implement all of these methods in MATLAB and set their parameters as recommended by the respective researchers. For the LBP operator, we apply the uniform LBP with 8 pixels in a circle of radius of 2 [7]. We use the same Gaussian filter of size 7 × 7 with the standard deviation of 1 to smooth the image for three methods, namely, our proposed ELDP, Weberface and gradientface methods.

Fig. 5 shows sample images from the Yale B face database and their corresponding preprocessed face images produced by the seven aforementioned compared techniques and the proposed ELDP code scheme. We use principle component analysis (PCA) with l2 norm as the classifier. In addition, we use the region-based histogram followed by χ2 dissimilarity measure to report the performance since this metric is used to compare the encoded features in [12]. However, our experiments show that PCA uses all eigenvectors followed by l2 yield better performance.

3.2 Results on PIE face database

The PIE database contains 41 368 grey-scale images (486 × 640 pixels) of 68 individuals. They are captured under various poses, illuminations and expressions. In our experiment, we use frontal images from the illumination subset (C27) that contains 21 images per subject. Fig. 4a shows all 21 images for a subject from this database. We conduct 21 experiments to evaluate the performance of the proposed method. In the ith experiment, we use the image i from each subject as the reference image and all the other 20 images as the test images. Fig. 6 shows the face recognition accuracy of different methods under each reference set using the PCA-based l2 measure. Obviously, the proposed ELDP method outperforms other methods for all reference images except the reference image 6. For this reference image, the gradientface method achieves slightly better accuracy. Table 1 summarises the average recognition accuracy of the eight methods for all reference images. Our proposed ELDP method achieves the highest average recognition rate of 98.29%, whereas the second highest rate is 96.63%, which is obtained by the gradientface method. Compared with the four methods in the same category, the proposed ELDP improves the face recognition accuracy of LBP, LDP, EnLDP and LDN by 3.06, 11.15, 4.78 and 9.70%, respectively. It clearly shows the effectiveness of the proposed ELDP code.
Fig. 7 shows the recognition accuracy of the eight methods under each reference set using the region-based \( \chi^2 \) measure. The recognition performance decreases for all methods when compared to the PCA-based \( l_2 \) measure.

3.3 Results on Yale B face database

The Yale B database contains grey-scale face images of ten individuals under nine poses. However, we use frontal face images in our experiment. Each subject has 64 images as well as an ambient image. The ambient image is an image with background illumination that is captured without a strobe going off. A sample ambient image is shown in the last row of Fig. 5. These frontal face images are categorised into six subsets based on the angle between the light source directions and the central camera axis: Subset 0 (0°, 60 images), Subset 1 (1°–12°, 80 images), Subset 2 (13°–25°, 100 images), Subset 3 (26°–50°, 120 images), Subset 4 (51°–77°, 100 images), Subset 5 (above 78°, 180 images) and Subset 5′ (above 78°+ ambient image, 190 images). Subset 0 is used for training. That is, 6 images per person are used for training. These 64 illumination conditions and the ambient image make the Yale B database more challenging to process than the PIE database.

Table 2 summarises the recognition rates of each of six subsets in the Yale B database for the eight compared methods using the PCA-based \( l_2 \) measure. We also include the average accuracy of subsets 1, 2, 3, 4 and 5′ in the last column since subset 5′ is a subset of subset 5′. The proposed ELDP achieves 100% recognition accuracy for all subsets, which is the best performance. Compared with the four methods in the same category, the proposed ELDP improves the face recognition accuracy of LBP, LDP, EnLDP and LDN by 1.73, 0.17, 1.20 and 1.54%, respectively. Comparing with the three methods in the other category, the proposed ELDP improves the face recognition accuracy of AdaS, gradientface and Weberface by 1.73, 0.17 and 0.17%, respectively.

Table 3 compares the performance of the eight methods using the region-based histogram-based \( \chi^2 \) measure. It clearly shows that the performance of all methods decreases especially for the AdaS method. However, ELDP still achieves the best average recognition accuracy as shown in the last column.

3.4 Results of the evaluation criterion

In this subsection, we compute an evaluation criterion, which offers an across comparison of average between-subjects and within-subjects, for each compared method. To this end, we first compute \( l_2 \) distances between all pairs of transformed images. For example, for PIE images with 68 subjects and 21 images per subject, we obtain a square matrix of size 1428 × 1428, which contains the pairwise \( l_2 \) distances between each image pair. For Yale B images with 10 subjects and 65 images per subject, we obtain a square matrix of size 650 × 650, which contains the pairwise \( l_2 \) distances between each image pair. So, for a face database with \( M \) subjects and \( K \) images per subject, this square matrix can be represented as

\[
\text{Mat}_{l_2} = [d_{i,j}]_{(M \times K) \times (M \times K)}
\]

where \( d_{i,j} \) represents the \( l_2 \) distance between the \( m \)th image of the subject \( i \) and the \( n \)th image of the subject \( j \). It should be emphasised that \( l_2 \) distance is computed between each ELDP image pair for the proposed method and is computed...
between each transformed image pair for the corresponding compared method.

We then use Mat\textsubscript{l} to compute another square matrix AveDiffMat\textsubscript{l} = [d′\textsubscript{ij}]\textsubscript{M×M}, which stores the average \textsubscript{l}2 distances between images of each subject and the average \textsubscript{l}2 distances between images of one subject and images of another subject. d′\textsubscript{ij} represents the average \textsubscript{l}2 distances between images of subject i and images of subject j and is computed as follows

\[
\begin{align*}
    d'_{ij} &= \frac{\sum_{m=1}^{K} \sum_{n=1}^{K} d_{mn}\textsubscript{ij}}{K \times K} & \text{if } i \neq j \\
    &= \frac{\sum_{m=1}^{K} \sum_{n=1}^{K} d_{mn}\textsubscript{ij}}{(K-1) \times (K-1)} & \text{if } i = j
\end{align*}
\]

Since the diagonal values (i.e. d\textsubscript{mm}\textsubscript{ij}) in Mat\textsubscript{l} represent distances from an image to itself, their values are 0’s and are removed from the calculation of d′\textsubscript{ij}.

We also use Mat\textsubscript{l} to compute the standard deviation matrix \(\varphi = [\sigma_{ij}]\textsubscript{M×M}\) where \(\sigma_{ij}\) is the corresponding standard deviation value for all the \(d_{mn}\textsubscript{ij}\) values used to compute d′\textsubscript{ij}.

We then compute the within-subject matrix \(W = [w_i]\textsubscript{i=1}^M\) and the between-subject matrix \(B = [b_i]\textsubscript{i=1}^M\) as well as their respective standard deviation matrices, \(\varphi_W = [\sigma_W^i]\textsubscript{i=1}^M\) and \(\varphi_B = [\sigma_B^i]\textsubscript{i=1}^M\) where

\[
    w_i = d'_{ii}, \quad b_i = \sum_{m\neq i}^{M} d'_{mi} / M - 1, \quad \sigma_W^i = \sigma_{ij} \quad \text{and} \quad \sigma_B^i = \frac{\sum_{m\neq i}^{M} \sigma_{ij}}{M-1}
\]

Finally, we compute an evaluation criterion, which offers an across comparison of average between-subjects and within-subjects criterion (BWSC), for each compared method by

\[
    BWSC_{\text{\textsubscript{l}}} = \frac{\sum_{i=1}^{M} [(b_i - w_i) - \max(\sigma_W^i, \sigma_B^i)]}{M}
\]
BWSC\textsubscript{L} is likely to achieve a better face recognition accuracy than the other methods.

Table 4 lists the values of BWSC\textsubscript{L} for each compared method when running on PIE and YaleB databases. It clearly shows that our proposed ELDP method achieves the highest values of BWSC\textsubscript{L} and therefore is more robust than the other methods. Our experimental results shown in Figs. 6 and 7 and Tables 1–3 further validate the effectiveness of this criterion. It is more interesting to observe that the proposed ELDP method and other compared methods achieve better performance on Yale B images, more challenging images, than on PIE images. This is mainly because that we use six images per subject for training for the Yale B database (the same setting has been used in the other peer methods) and use one image per subject for training for the PIE database. However, our
Table 1  Average recognition accuracy (%) for PIE face images using the PCA-based $l_2$ measure

<table>
<thead>
<tr>
<th>Method</th>
<th>AdaS</th>
<th>Gradientface</th>
<th>Weberface</th>
<th>LBP</th>
<th>LDP</th>
<th>EnLDP</th>
<th>LDN</th>
<th>ELDP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>89.63</td>
<td>96.63</td>
<td>94.55</td>
<td>95.37</td>
<td>88.43</td>
<td>93.81</td>
<td>89.60</td>
<td>98.29</td>
</tr>
</tbody>
</table>

Table 2  Recognition accuracy (%) for Yale B face images using PCA-based $l_2$ measure

<table>
<thead>
<tr>
<th>Method</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S5'</th>
<th>Ave.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AdaS</td>
<td>100</td>
<td>100</td>
<td>99.17</td>
<td>100</td>
<td>99.44</td>
<td>95.26</td>
<td>98.30</td>
</tr>
<tr>
<td>Gradientface</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Weberface</td>
<td>100</td>
<td>100</td>
<td>97.17</td>
<td>100</td>
<td>99.44</td>
<td>95.26</td>
<td>98.30</td>
</tr>
<tr>
<td>LBP</td>
<td>100</td>
<td>100</td>
<td>97.17</td>
<td>100</td>
<td>99.44</td>
<td>95.26</td>
<td>98.30</td>
</tr>
<tr>
<td>LDP</td>
<td>100</td>
<td>100</td>
<td>97.17</td>
<td>100</td>
<td>99.44</td>
<td>95.26</td>
<td>98.30</td>
</tr>
<tr>
<td>EnLDP</td>
<td>100</td>
<td>100</td>
<td>97.17</td>
<td>100</td>
<td>99.44</td>
<td>95.26</td>
<td>98.30</td>
</tr>
<tr>
<td>LDN</td>
<td>100</td>
<td>100</td>
<td>97.17</td>
<td>100</td>
<td>99.44</td>
<td>95.26</td>
<td>98.30</td>
</tr>
<tr>
<td>ELDP</td>
<td>100</td>
<td>100</td>
<td>97.17</td>
<td>100</td>
<td>99.44</td>
<td>95.26</td>
<td>98.30</td>
</tr>
</tbody>
</table>

Table 3  Recognition accuracy (%) for the Yale B face images using region-based histogram-based $\chi^2$ measure

<table>
<thead>
<tr>
<th>Method</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S5'</th>
<th>Ave.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AdaS</td>
<td>100</td>
<td>100</td>
<td>89.17</td>
<td>60</td>
<td>46.11</td>
<td>45.26</td>
<td>73.39</td>
</tr>
<tr>
<td>Gradientface</td>
<td>100</td>
<td>100</td>
<td>98.33</td>
<td>97</td>
<td>96.67</td>
<td>92.63</td>
<td>96.78</td>
</tr>
<tr>
<td>Weberface</td>
<td>100</td>
<td>100</td>
<td>96.67</td>
<td>96</td>
<td>92.22</td>
<td>87.37</td>
<td>94.58</td>
</tr>
<tr>
<td>LBP</td>
<td>100</td>
<td>100</td>
<td>96.67</td>
<td>95</td>
<td>87.22</td>
<td>83.16</td>
<td>93.39</td>
</tr>
<tr>
<td>LDP</td>
<td>100</td>
<td>100</td>
<td>97.17</td>
<td>96</td>
<td>94.44</td>
<td>93.16</td>
<td>96.95</td>
</tr>
<tr>
<td>EnLDP</td>
<td>100</td>
<td>100</td>
<td>97.17</td>
<td>97</td>
<td>97.22</td>
<td>97.37</td>
<td>98.81</td>
</tr>
<tr>
<td>LDN</td>
<td>100</td>
<td>100</td>
<td>97.17</td>
<td>97</td>
<td>96.67</td>
<td>95.26</td>
<td>97.63</td>
</tr>
<tr>
<td>ELDP</td>
<td>100</td>
<td>100</td>
<td>97.17</td>
<td>95</td>
<td>97.22</td>
<td>96.32</td>
<td>97.80</td>
</tr>
</tbody>
</table>

Table 4  BWSC values computed by (2) for both PIE and Yale B face images

<table>
<thead>
<tr>
<th>Method</th>
<th>AdaS</th>
<th>Gradientface</th>
<th>Weberface</th>
<th>LBP</th>
<th>LDP</th>
<th>EnLDP</th>
<th>LDN</th>
<th>ELDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIE</td>
<td>1.464724</td>
<td>3.812303</td>
<td>5.729437</td>
<td>1.609345</td>
<td>0.712541</td>
<td>3.340555</td>
<td>2.130051</td>
<td>8.520582</td>
</tr>
<tr>
<td>Yale B</td>
<td>0.937955</td>
<td>0.189013</td>
<td>0.884916</td>
<td>0.294389</td>
<td>-0.67606</td>
<td>-0.68287</td>
<td>-1.16753</td>
<td>1.051299</td>
</tr>
</tbody>
</table>

Fig. 7  Comparison of recognition accuracy for PIE face images using the region-based histogram-based $\chi^2$ measure

IET Biom., pp. 1–8
results shown in Table 4 indicate each compared method has a smaller evaluation value for the Yale B database than for the PIE database. In other words, our evaluation criterion also verifies the Yale B database is more challenging than the PIE database.

4 Conclusions

We propose a new illumination invariant face descriptor, ELDP, to preprocess face images with illumination variations. The proposed ELDP method uses Kirsch compass masks to compute eight directional edge responses in the gradient domain, which is robust against illumination changes and noise. It extracts more edge information by utilizing all eight directional edge images. It also indicates valuable structural information from the neighbourhood by considering the relations among eight directional edge responses. Our experiments on two face databases (PIE and Yale) illustrate the effectiveness of the proposed ELDP method and the proposed evaluation criterion and demonstrate the proposed scheme achieves the best face recognition accuracy when compared to seven state-of-the-art methods.

5 References

Author Queries
Mohammad Reza Faraji, Xiaojun Qi

Q1  Please retain the email address of the corresponding author only.
Q2  Please provide issue number.