Face recognition based on stable uniform patterns

A. Mallikarjuna Reddy 1 *, V. Venkata Krishna 2, L. Sumalatha 3

1 Research Scholar, JNTU Kakinada, Asst. Professor, Anurag Group of Institutions (Autonomous), Hyderabad, India
2 Principal, Nishitha College of Engineering, Hyderabad, India
3 Professor, University College of Engineering, JNTUK, Kakinada, Andhra Pradesh, India
*Corresponding author E-mail: mallikarjunreddycse@csvr.ac.in

Abstract

Face recognition (FR) is one of the challenging and active research fields of image processing, computer vision and biometrics with numerous proposed systems. We present a feature extraction method named “stable uniform local pattern (SULP)”, a refined variant of ULBP operator, for robust face recognition. The SULP directly applied on gradient face images (in x and y directions) of a single image for capturing significant fundamental local texture patterns to build up a feature vector of a face image. Histogram sequences of SULP images of the two gradient images are finally concatenated to form the “stable uniform local pattern gradient (SULPG)” vector for the given image. The SULPG approach is experimented on Yale, ATT-ORL, FERET, CAS-PEAL and LFW face databases and the results are compared with the LBP model and various variants of LBP descriptor. The results indicate that the present descriptor is more powerful against a wide range of challenges, such as illumination, expression and pose variations and outperforms the state-of-the-art methods based on LBP.

Keywords: Expressions; Gradient; Histogram; Illumination; Pose; Robust; Uniform.

1. Introduction

Face recognition is one of the significant and challenging topics for decades in computer vision and pattern recognition and [1], [2]. This is due to its widespread practical applications, such as access control [3], smart card design [4], information security [5], surveillance system [6], expression cloning [7] and law enforcement follow-up [8]. Facial images captured under the real world environment may have variations, dissimilar poses, varying illumination and occlusions such as sunglasses etc... These issues make the face recognition system as a challenging and interesting problem. Numerous FR methods are proposed in the literature to tackle the above challenges [9], [10]. There are three most popular and dominant categories of face recognition systems i.e. the methods based on deep learning [11], holistic (global features) approaches [12], and face image descriptors (local feature or component based) [13], [14]. The well-known learning-based descriptors for face recognition are Local Quantized Patterns (LQP) [15], and Discriminate Face Descriptor (DFD) [16]. The performance of these methods [15, 16] depends on unsupervised or supervised learning techniques. In the holistic method global features are extracted from the entire face image. The local feature based methods decomposes face image into local facial features. These local features represent the most discriminant intrinsic information of face image to build feature vector. The feature extraction methods should be robust against many challenges. The face image descriptor methods have several advantages like ease of use, data independence, and robustness to real-life challenges such as illumination and expression differences. The LBP [17] and Gabor wavelets [18] based methods are most well-known and popular methods for feature based face recognition. The local binary pattern (LBP) [19, 20] have aroused increasing interest among the researchers in image processing, pattern recognition and computer vision, due to its simplicity, tolerance to monotonic illumination changes and efficient representation of local structures of images. Researchers worked on LBP and derived facial features efficiently for different tasks, which include face recognition [21], [22], [23], face detection [24], facial expression analysis [26], [26], demographic (gender, race, age, etc.) classification [27]. Many variants of LBP have been proposed in the literature to improve its efficiency further and LBP based methods are combined with statistical methods. The other well-known variants of LBP include Transition LBP (TLBP) [28] and Direction coded LBP (DLBP) [28]. The methods based on Gabor wavelets [29] are proposed in the recent literature, to obtain more precisely multi-scale and multi-orientation information of face images. A method for recognition on blurred faces [30] and another method based on edge magnitudes [31] are also proposed in the literature.

The existing face image descriptors, achieved successful performance but at the cost of high dimensionality and computational cost. Therefore there is a necessity to derive a face descriptor with superior performance with lower computational cost and feature size. To address this, in this work, we propose a novel feature extraction method called as Stable Uniform Local Pattern of gradients (SULPG) for FR. This descriptor provides an integrated way to capture stable uniform local patterns from two gradient images corresponding to a single image.

The remainder of the paper is organized as follows: Section 2 details the proposed SULPG descriptor. Experimental results are presented in Section 3, leading to conclusions in Section 4.

2. Formulation of stable uniform local pattern (SULP)

An image is composed of micro patterns. The LBP can be viewed as first-order circular derivative patterns. The derivation of LBP...
code is a “two-step” process. In the initial step, each neighboring pixel of a “3 x 3 neighborhood”, is compared with its central pixel and binary patterns are derived using equation 1. In the second step, a unique decimal number, which is called as LBP or ULBP code, is derived based on equation 2. The LBP code holds significant information about the distribution of edges and other local features in an image. The neighboring pixels are denoted by $p = \{p_0, p_1, \ldots, p_9\}$, the central pixel by $p_c$ and $R$ represents the radius of the neighborhood.

$$b_i = \begin{cases} 1 & \Delta p_i \geq 0 \\ 0 & \Delta p_i < 0 \end{cases}$$

(1)

Where $\Delta p_i = p_i - p_c$.

For each 3x3 neighborhood, a unique LBP code is derived from the equation (2)

$$\text{LBP}_{PB} = \sum_{i=0}^{8} b_i \times 2^i$$

(2)

The LBP$_{PB}$ operator produces $2^P$ different output values.

The basic LBP operator was later derived for different number of neighboring pixels $P$ and radius $R$. The uniform patterns are introduced by mannepaa et al. [32]. It is observed that more than 90% of the windows in human faces and natural textures are uniform LBP (ULBP). This proposition was further confirmed by ojal et al. [19]. The commonality in uniform LBP (ULBPs) is that they contain at most two, “zero to one or one to zero transitions” in the circular binary code. The uniformity of “U” of a neighborhood $G$, is measured in the following way

$$U(G) = \text{abs} \left( \sum_{g=1}^{r-1} g_{g-1}^{r-1} - g_{g+1}^{r-1} \right) + \sum_{g=1}^{r-1} S(g_{g-1}^{r-1}) - S(g_{g+1}^{r-1})$$

(3)

The neighborhood with $U(G)$ is less than or equal to two are termed as uniform local binary pattern (ULBP). The total number of uniform patterns on a circular neighborhood of $(P, R)$ is $P + (P-1) + 2$. The remaining $2^P - (P + (P-1) + 2)$ codes are termed as non-ULBP (NULBP) codes. However, it is noted that at most 10% of the windows in the texture images will have NULBP windows. It is observed that NULBPs generate noise and high dimensionality. Therefore the NULBPs are represented under one label and treated as miscellaneous. Allocating one label for all large number of NULBPs drastically reduces overall dimensionality and also overcomes the problem of noise patterns. There are 58 ULBPs and 198 NULBPs on a 3x3 window with $P=8$ and $R=1$. In this case, the dimensionality of the histogram by considering the LBP and ULBP will be $256$ and $59$ (in this case all NULBPs are treated under one label). The present paper gained advantage in deriving efficient feature vectors from facial images by initially computing gradient images because gradients are proved to be more sensitive of visual human system than local intensity differences. Extraction of texture information from gradient images is more efficient than raw image [33, 34]. Then, on these directional gradient images the present paper derives SULP feature vectors to extract stable uniform local patterns. The stable Uniform local patterns which are binary strings derived based on bitwise transitions from 0 to 1 or 1 to 0, in a circular manner are utilized to reduce the SULPG description’s length. The SULP histogram sequences of these two gradient images (SULP-G and SULP-G$_{o}$ a single image) are concatenated to derive SULPG description of the image. This makes the feature size (number of histogram bins) of SULPG as 2 x 43 = 86.

### 3. Results and discussions

In this section, the performance of SULPG descriptor is computed. To test the advantages of SULPG, the performance of seven state-of-the-art face image descriptors, i.e., LBP [20], LTP [22], ULBP[32], LPQ [30], PLBP [35], multi-scale LBP (MsLBP) [19], Multi-scale DLBP (MsDLBP) [36], are also presented. These existing descriptors [19], [20], [22], [30], [32] are easy to implement and they are based on local features. The cropped face image data for these descriptors in the present paper are exactly the same. For efficient face recognition the present paper evaluated histograms of LBP, ULBP, LTP, MPLBP and SPLBP on each individual facial image and placed in training database. In the similar way the above histograms are evaluated for test facial image and the face recognition is evaluated based on Chi-square distance metric as given in equation 6.

$$R(d, t) = \min (\sum_{i=1}^{n} (d_i - t_i)^2/(d_i + t_i))/2$$

(6)

Where $d$, $t$ are two image features (histogram vectors) and $R (d, t)$ is the histogram distance for recognition.
For experimental sake this paper used the most popular and widely used facial databases, i.e., FERET (face recognition technology) [37], CAS-PEAL [38], LFW (Labeled Faces in the Wild) [39]; Yale Face Database B [40]; AT&T ORL [41].

FERET database contains two sub categories of databases namely, Frontal FERET Image Sets and Non-Frontal FERET Image Sets. The frontal image sets i.e. F1, F2, Duplicate I (Dup I) and Duplicate II (Dup II) are part of the Frontal FERET Image Set. There F1 set, images are used for gallery and training sets. The F4, F5, Dup I and Dup II image sets are used for probe sets and they consist of 1195, 194, 722 and 234 images respectively. Apart from frontal images, FERET has pose-view images of 200 people, which are Non-Frontal FERET Image Sets. This paper considered only the frontal FERET data sets for our experimental purpose.

The CAS-PEAL face database includes Chinese face database with different sources of variations, especially Pose, Expression, Accessories, and Lighting (PEAL). Currently, the CAS-PEAL face database contains 99,594 images of 1040 individuals (595 males and 445 females) with varying Pose, Expression, Accessory, and Lighting (PEAL). A subset of the whole CAS-PEAL face database referred as CAS-PEAL-R1, with 30,900 images of 1040 subjects is publicly available now. The CAS-PEAL-R1 images are divided into two categories: the frontal and non-frontal. This database contains one training set, one gallery set, and six probe sets in frontal category and three in non-frontal category. The six probe sets of frontal category are denoted as PE, PA, PL, PT, PB, and PS correspond to variations in expression, accessory, lighting, time, background, and the distance of frontal faces, respectively. The three categories of non-frontal probe sets PU, PM, and PD correspond to one type of pose variation i.e. looking: upward, towards camera and downwards.

The LFW [39], face photographs are designed for studying the problem of unconstrained face recognition. The data set contains more than 13,000 images of faces organized in two views: view1 and view 2. LFW is a challenging dataset since its images contain large pose, expression, and illumination variations. For model identification and parameter tuning the view 1 set of LFW is used. For performance reporting View 2 is used. In View 2 two models are used to make use of the training set: image-restricted paradigm and unrestricted paradigm. In the first model the training is made on officially defined images. In the second model the identity information of the training images can be used. For model identification and parameter tuning the view 1 set of LFW is used. There are 5760 images of 10 subjects are available in the Yale Face Database B. These images are derived from 576 viewing conditions (9 poses x 64 illumination conditions).

The sample images of these databases are shown in Fig. 1, 2, 3 and 4. The SULP-G and SULP-G facial images of the present descriptor are shown in Fig. 5, 6, 7 and 8 for facial data bases YALE, FERET, CAS-PEAL and LFW respectively.
Fig. 3: Sample facial images of CAS-PEAL database.

Fig. 4: Sample Facial Images of LFW Face Database.
Fig. 5: (A) Original YALE Image (B) X-Gradient Image (C) Y-Gradient Image (D) SULP-Gximage (E) SULP-Gyimage

Fig. 6: (A) Original FERET Image (B) X-Gradient Image (C) Y-Gradient Image (D) SULP-Gximage (E) SULP-Gyimage.
Fig. 7: (A) Original CAS-PÉAL Image (B) X-Gradient Image (C) Y-Gradient Image (D) SULP-Gximage (E) SULP-Gyimage.

Fig. 8: (A) Original LFW Image (B) X-Gradient Image (C) Y-Gradient Image (D) SULP-Gximage (E) SULP-Gyimage.
The FR rate results on FERET frontal datasets are recorded in Table 3 and the following observations are made:

1) High performance is recorded by all descriptors on Fb and Fc probe sets when compared to DUP1 and DUP2 probe sets. This is due to the well-controlled environment of Fb and FC probe sets. The SULPG achieves more than 1% of high mean classification rate when compared to other descriptors. The descriptors exhibited a little high recognition rate on probe set Fc when compared to Fb.

2) There is a considerable drop in the performance of DUP1 and DUP2 probe sets by all descriptors. This is because the images in DUP1 and DUP2 exhibit expression and illumination variation when compared to Fb and Fc.

3) The proposed SULPG has shown high recognition rate when compared to other descriptors at low dimensionality.

4) The present paper also computed the overall mean recognition rate for Fb and Fc probe sets and also for DUP1 and DUP2 probe sets and listed Table 2.

The FR rate results on FERET frontal datasets are recorded in Table 3 and the following observations are made:

1) All descriptors have shown high FR rate on frontal probe subsets i.e. PE, PA, PT, PB, PS. The proposed SULPG resulted with the high FR rate on all these five probe sets.

2) All descriptors except LPQ have shown a poor FR rate around 45% on the PL probe set. This is mainly because of high illumination and image blur variation of PL set. The LPQ could sustain from these variations and resulted a 10% of the high FR rate i.e. around 57%.

3) We have also computed the mean recognition rate of frontal subsets (excluding and also including PL) and listed in the Table 3. The mean FR rate of SULP is higher than the rest of the descriptors with low dimensionality except LPQ.

4) On non-frontal probe sets, the probe set PM has exhibited high performance due to the pose of the face with respect to camera. After PM the probe set PU attained better performance than PD because capturing of facial information will be more in this pose (PU) when compared to PD.

5) The mean FR rate clearly exhibits the superiority of the present descriptor and robustness to pose variations on probe sets PM, PD, PU over the rest of the descriptors.

The recognition rate on LFW dataset is reported in Table 4 and the following are note down.

Table 2: Identification Rates for Different Descriptors on FERET.

<table>
<thead>
<tr>
<th>Descriptors</th>
<th>PE</th>
<th>PA</th>
<th>PT</th>
<th>PB</th>
<th>PS</th>
<th>PL</th>
<th>PU</th>
<th>PM</th>
<th>PD</th>
<th>Mean (PE-PL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>94.27</td>
<td>91.82</td>
<td>100.0</td>
<td>99.46</td>
<td>99.64</td>
<td>46.90</td>
<td>60.32</td>
<td>83.66</td>
<td>44.60</td>
<td>97.04</td>
</tr>
<tr>
<td>LTP</td>
<td>94.39</td>
<td>91.77</td>
<td>100.0</td>
<td>99.46</td>
<td>99.64</td>
<td>47.17</td>
<td>61.32</td>
<td>84.46</td>
<td>44.68</td>
<td>97.05</td>
</tr>
<tr>
<td>LPQ</td>
<td>93.95</td>
<td>92.39</td>
<td>100.0</td>
<td>99.28</td>
<td>99.27</td>
<td>57.16</td>
<td>65.98</td>
<td>86.46</td>
<td>44.76</td>
<td>96.98</td>
</tr>
<tr>
<td>PLBP</td>
<td>94.07</td>
<td>91.92</td>
<td>100.0</td>
<td>99.28</td>
<td>99.48</td>
<td>45.24</td>
<td>58.28</td>
<td>82.78</td>
<td>43.92</td>
<td>96.95</td>
</tr>
<tr>
<td>MsLBP</td>
<td>94.87</td>
<td>92.02</td>
<td>100.0</td>
<td>99.86</td>
<td>99.84</td>
<td>46.02</td>
<td>60.01</td>
<td>82.94</td>
<td>44.28</td>
<td>97.32</td>
</tr>
<tr>
<td>MsDLBP</td>
<td>95.16</td>
<td>92.04</td>
<td>100.0</td>
<td>99.46</td>
<td>99.64</td>
<td>47.75</td>
<td>61.76</td>
<td>85.16</td>
<td>44.88</td>
<td>97.26</td>
</tr>
<tr>
<td>SULPG</td>
<td>95.46</td>
<td>93.72</td>
<td>100.0</td>
<td>99.48</td>
<td>99.27</td>
<td>48.11</td>
<td>55.00</td>
<td>82.03</td>
<td>37.03</td>
<td>96.32</td>
</tr>
</tbody>
</table>

Table 3: Rank-1 Identification Rates for Different Face Image Descriptors on the Nine Probe Sets of PEAL.

<table>
<thead>
<tr>
<th>Descriptors</th>
<th>PE</th>
<th>PA</th>
<th>PT</th>
<th>PB</th>
<th>PS</th>
<th>PL</th>
<th>PU</th>
<th>PM</th>
<th>PD</th>
<th>Mean (PE-PL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>94.27</td>
<td>91.82</td>
<td>100.0</td>
<td>99.46</td>
<td>99.64</td>
<td>46.90</td>
<td>60.32</td>
<td>83.66</td>
<td>44.60</td>
<td>97.04</td>
</tr>
<tr>
<td>LTP</td>
<td>94.39</td>
<td>91.77</td>
<td>100.0</td>
<td>99.46</td>
<td>99.64</td>
<td>47.17</td>
<td>61.32</td>
<td>84.46</td>
<td>44.68</td>
<td>97.05</td>
</tr>
<tr>
<td>LPQ</td>
<td>93.95</td>
<td>92.39</td>
<td>100.0</td>
<td>99.28</td>
<td>99.27</td>
<td>57.16</td>
<td>65.98</td>
<td>86.46</td>
<td>44.76</td>
<td>96.98</td>
</tr>
<tr>
<td>PLBP</td>
<td>94.07</td>
<td>91.92</td>
<td>100.0</td>
<td>99.28</td>
<td>99.48</td>
<td>45.24</td>
<td>58.28</td>
<td>82.78</td>
<td>43.92</td>
<td>96.95</td>
</tr>
<tr>
<td>MsLBP</td>
<td>94.87</td>
<td>92.02</td>
<td>100.0</td>
<td>99.86</td>
<td>99.84</td>
<td>46.02</td>
<td>60.01</td>
<td>82.94</td>
<td>44.28</td>
<td>97.32</td>
</tr>
<tr>
<td>MsDLBP</td>
<td>95.16</td>
<td>92.04</td>
<td>100.0</td>
<td>99.46</td>
<td>99.64</td>
<td>47.75</td>
<td>61.76</td>
<td>85.16</td>
<td>44.88</td>
<td>97.26</td>
</tr>
<tr>
<td>SULPG</td>
<td>95.46</td>
<td>93.72</td>
<td>100.0</td>
<td>99.48</td>
<td>99.27</td>
<td>48.11</td>
<td>55.00</td>
<td>82.03</td>
<td>37.03</td>
<td>96.32</td>
</tr>
</tbody>
</table>

The FR rates of Yale face database B are listed in the Table 5 and the following are noted.

1) All the descriptors achieved the FR rate around 75% for LF database set and this is due to large variations in pose, expression, and illumination in LF.
2) The proposed SULPG achieved superior performance on LF database set when compared to other descriptors.

The FR rates of Yale face database B are listed in the Table 5 and the following are noted.

1) All the descriptors has shown high FR rate and this may be due to small dataset.
2) The LPQ has exhibited a high recognition rate because LPQ is viewed as illumination invariant.

Table 4: Recognition Rate on LFW Database.

<table>
<thead>
<tr>
<th>Descriptors</th>
<th>LFW</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>72.43</td>
</tr>
<tr>
<td>LTP</td>
<td>72.65</td>
</tr>
<tr>
<td>ULBP</td>
<td>73.27</td>
</tr>
<tr>
<td>LPQ</td>
<td>72.68</td>
</tr>
<tr>
<td>PLBP</td>
<td>72.78</td>
</tr>
<tr>
<td>MsLBP</td>
<td>72.88</td>
</tr>
<tr>
<td>MsDLBP</td>
<td>72.17</td>
</tr>
<tr>
<td>SULPG</td>
<td>73.54</td>
</tr>
</tbody>
</table>

Table 5: Recognition Rate on the YALE-B Database.

<table>
<thead>
<tr>
<th>Descriptors</th>
<th>YALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>81.94</td>
</tr>
<tr>
<td>LTP</td>
<td>82.01</td>
</tr>
<tr>
<td>ULBP</td>
<td>83.61</td>
</tr>
<tr>
<td>LPQ</td>
<td>82.43</td>
</tr>
<tr>
<td>PLBP</td>
<td>83.06</td>
</tr>
<tr>
<td>MsLBP</td>
<td>83.86</td>
</tr>
<tr>
<td>MsDLBP</td>
<td>81.97</td>
</tr>
<tr>
<td>SULPG</td>
<td>85.78</td>
</tr>
</tbody>
</table>
The present paper computed the feature vector size the (number of histogram bins or dimension) for each descriptor and listed Table 6.

Table 6: Feature Sizes of the Proposed and Existing Facial Image Descriptors

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Feature size</th>
<th>Descriptor</th>
<th>Feature size</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTP[22]</td>
<td>512</td>
<td>MsLBP[19]</td>
<td>512</td>
</tr>
<tr>
<td>LPQ[30]</td>
<td>256</td>
<td>Proposed-SULPG</td>
<td>86 (2 x 43)</td>
</tr>
<tr>
<td>PULBP</td>
<td>96</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Table 8 clearly demonstrates that the feature vector dimensions of LBP and LPQ are almost three times higher than SULPG, whereas the LTP, MsLBP and MsDLBP feature vectors are six times higher than SULPG. Finally, from this and from the above experimental results, it clearly evident that SULPG with low computational cost than other descriptors, has shown high discriminative power and excellent robustness to various variations of human faces i.e. pose, expression, moderate illumination variation, accessory and lighting. The performance of the SULPG is not affected wherever the images exhibits high illumination and blur variations. This is mainly to gradient strategy on SULPG images. The SULPG is experimented on large size databases and the descriptor extracted more discriminative information in the form of stable transitions, which are subsets of fundamental units of texture. The excellent performance of SULPG can be explained as follows. It encodes the stable fundamental information of facial images while treating non-significant and noisy information of facial images under one label and the gradient strategy ensures good complementarily in dealing with illumination and blur variations.

4. Conclusions

In this paper we presented SULPG, a novel feature extraction method for Face identification, based on the impressive results of visual perception researches [33-34]. After an in-depth study, on the descriptors based on LBP in various classifications, recognition analysis issues, the present paper derived a new variant of LBP called SULPG with limited, consistent and stable features for face representation. The SULPG represents significant texture features like lines, curves, edges, holes, blobs etc... The SULPG descriptor is exploited on two gradient images i.e. on x and y axis of the raw image to generate SULPG image. The enhanced SULPG contains more discriminative information and it overcomes the illumination effect. The proposed SULPG descriptor is compared with other existing descriptors in a comprehensive and systematic manner and tested on five large facial databases. The method achieves stable and high performance over the other descriptors with respect to various challenges such as illumination, occlusion, facial expressions, pose variations and time-lapse images, by reducing the number of image features by reducing the overall dimensionality.

References


