<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Enhancing local binary patterns distinctiveness for face representation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Ghahramani, Mohammad; Yau, Wei-Yun; Teoh, Eam Khwang</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>2011</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/18009">http://hdl.handle.net/10220/18009</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© 2011 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: [<a href="http://dx.doi.org/10.1109/ISM.2011.78">http://dx.doi.org/10.1109/ISM.2011.78</a>].</td>
</tr>
</tbody>
</table>
Enhancing Local Binary Patterns Distinctiveness for Face Representation

M. Ghahramani*1
Dept. of Electrical and Electronics Engineering
Nanyang Technological University, NTU*1
Singapore, 639798
moha0116@ntu.edu.sg

W. Y. Yau#2 and E. K. Teoh*1
Institute for Infocomm Research, I2R#2
1 Fusionopolis Way.
Singapore, 138632
wyyau@i2r.a-star.edu.sg, eekteoh@ntu.edu.sg

Abstract—The Local Binary pattern (LBP) is a well-known feature and has been widely used for human identification. However, the amount of information extracted is limited which reduces the LBP discriminative power. Recently, some enhancements have been proposed by adding preprocessing stages or considering more neighbor pixels to enrich the extracted feature. In this paper, we propose Uniformly-sampled Thresholds for LBP (UTLBP) operator that increases the richness of information derived from the LBP feature. It outperforms other features in various probe sets of the large CAS-PEAL database for face recognition. Moreover, we collected a database of 25 families to verify the superiority of the proposed feature in the family verification. Results show that using the UTLBP, the total error in face recognition and family verification is reduced up to 8% and 3% respectively comparing to the state of the art LBP. It improves the missing family member verification performance up to 3% where, contrary to expectation, increasing the LBP operator radius worsens the performance by 2%.

Keywords- LBP modification; feature; family verification

I. INTRODUCTION

Nowadays, most of successful algorithms for face detection and matching are learning-based. These algorithms learn from the pool of extracted features to perform classification. The aim of feature design is to obtain discriminative, fast and low dimensional features. In the literature, several features are investigated to extract useful information of the face for further processing such as detection, classification and recognition [1]. Among them, Local Binary Patterns (LBP) is getting more popular for human and face detection/recognition since it is computationally inexpensive, simple and reliable [2, 3].

However, LBP operator extracts only limited amount of information from the image. For example, it cannot differentiate if a pixel is greater or equal to the center pixel of the operator's mask [4]. In this paper, our aim is to allow the LBP operator to extract more information from the image to enhance its discriminative power. We then evaluate the proposed modified LBP on recognition tasks of face recognition under various changing conditions and the challenging problem of “family verification” (or family classification), which is to recognize whether a group of people are related based on facial characteristics [5]. There are real life applications that could benefit from family verification as it has been reported on the average 2,185 missing children each day in the united states [6] and recently, a mother could find her missing abducted daughter in Facebook by her particulars which was a rare online success in the search for missing kids online as experts say [7].

The rest of the paper is organized as follows: Section 2 reviews the various modifications to LBP and similar operators in the literature. Section 3 then introduces the method to extract information missing in the LBP operator, followed by further analysis of the proposed enhanced operator. These operators are then compared by evaluating their performance when applied to face recognition in terms of robustness against background, lighting, aging, accessory and distance changes using the CAS-PEAL database. The proposed operator dominance is also checked in family verification in Section 4. Finally the conclusion deduces the whole paper in Section 5.

II. LITERATURE REVIEW

A. Local Binary Patterns (LBP)

The LBP operator [3] is one of the fast and reliable features for human detection and identification [8]. Ahonen et al. studied its robustness against monotonic gray level changes in [8]. Therefore, it is popular for demanding image analysis and classification problems such as face recognition.

The human being recognizes faces similarities by composing micro face patterns which can be described by the LBP features [8]. The original LBP operator assigns a binary code to every pixel of an image by thresholding the surrounding pixels around the center pixel. The sign function \( s(x) \) output is “1” if the input is greater or equal to zero and “0” if otherwise as in equation (1).

\[
\begin{align*}
s(x) &= \begin{cases} 
1, & x \geq 0 \\
0, & x < 0 
\end{cases}
\end{align*}
\] (1)

The LBP feature is then given by,

\[
LBP_{P,R}(c) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p
\] (2)
where \( g_c \) is the gray level of the center pixel and \( g_p \) is the gray level of the surrounding pixel.

Ojala et al. enhanced the LBP operator by considering more neighbors \( 'P' \) around the center pixel [3]. These neighbors are located at sampling points on the circle centered at the center pixel with radius \( 'R' \). They utilized bilinear interpolation to approximate the sampling point value if the sampling point falls close to the pixels borders.

The modified operator is denoted as \( \text{LBP}(P,R) \) where \( 'P' \) indicates the number of sampling points on the desired circle with radius \( 'R' \). Aside from changing parameters of LBP operator, the following two research works increased the amount of information extracted from the image by LBP in the literature.

B. Local Ternary Patterns (LTP)

According to function \( s(x) \) in the LBP operator, it fails to distinct between the patterns generated from neighboring pixels with equal value, and those with larger pixel levels. He and Cercone [4] managed to tackle this problem by theLocal Triplet Patterns (LTP) [4] while maintaining the LBP feature structure and capabilities, for image retrieval. They changed the function \( s(x) \) to return three distinctive results (smaller, equal and larger) when comparing pixel values as,

\[
LTP_{P,R}(c) = \sum_{p=0}^{P-1} s_{LTP}(g_p - g_c)3^p
\]  

where

\[
s_{LTP}(x) = \begin{cases} 
2, & x > 0 \\
1, & x = 0 \\
0, & x < 0 
\end{cases}
\]

Hence a sequence of triplet codes leads to a large and sparse histogram of feature values. To tackle this problem they proposed a scaling and a neighboring parameter to quantize the feature to a smaller limit which reduces the feature distinctiveness [4]. In their experiments different scales led to varying results in separate conditions such that a constant value could not be easily generalized. The experiments conducted showed the superiority of the LTP to the LBP operator in most of experiments.

C. Local Derivative Patterns (LDP)

Zhang et al. explored the feasibility and usefulness of using high-order local patterns for face representation. In this scheme, LBP is considered as the non-directional first-order local pattern operator since LBP determines all-direction first-order derivative binary result and misses more detailed information enclosed in the input object [9]. Local Derivative Pattern (LDP) was proposed to encode directional pattern features based on local derivative variations. It encodes the higher-order \( (n^{th}) \) derivative information which contains more detailed discriminative features that the first-order local pattern (LBP) fails to extract from an image as claimed by the authors [9].

III. THE PROPOSED OPERATOR

A closer and deeper look at the LBP operator stated in equation (2) shows that, there are two types of information that exist at the right part of the equation but vanish at the left side of equation (2); the pixel value of the center pixel \( g_c \), and the sign and magnitude of difference between the neighboring pixels with the center pixel, \( (g_p - g_c) \). The pixel value information \( g_c \) is sensitive to variations such as lighting or background and that is why we design features to make the classification/recognition robust to these changes.

The conventional LBP operator quantizes the relative intensity information of the surrounding pixels to the center pixel. In this way, large variations of the intensity level do not change the LBP feature considerably since the relative gray scale information of neighbor pixels is almost unchanged. However, the quantization error of the conventional LBP feature is very high due to the step function \( s(x) \) with one level quantization at “0” level. We thus propose to include the missing information using Uniformly-sampled Thresholds for LBP (UTLBP) as explained below. Note that, the pixel depth \( 'PixD' \) is “255” for our analysis below although other pixel depth values are applicable.

A. Uniformly-sampled Thresholds for Local Binary Pattern (UTLBP)

The pure relative information of the surrounding pixel value to the center pixel can be embedded into the LBP operator by the term \( 'g_p - g_c' \). If we omit the step function to keep the pure grayscale information \( 'g_p - g_c' \) the feature evaluated for a 3×3 window would be as large as \( 2×256^6 \) or approximately \( 2×10^{20} \), (2 is multiplied due to the sign bit). To tackle this problem, we consider Uniformly-sampled Thresholds for LBP (UTLBP) to evaluate the binary code for each threshold. Thresholds are uniformly distributed within the range \( (-PixD<Thr<PixD) \). The feature vector of pixel ‘c’ is modeled as,

\[
\text{UTLBP}_{P,R}(c) = [\text{utlbp}_{P,R}(c,Th_{T1}), ..., \text{utlbp}_{P,R}(c,Th_{T})]
\]

where,

\[
\text{utlbp}_{P,R}(c,Th_k) = \sum_{p=0}^{P-1} s(g_p - g_c - Th_k)2^p
\]

Each feature subset is extracted by employing \( 'Th_{Tj}' \) where \( 1\leq k \leq T \). Note that, increasing the number of LBP neighborhood pixels from \( P \) to \( Q \) [8], enlarges the feature dimension by order of \( 2^{(Q-P)} \) while increasing number of thresholds from \( T \) to \( T_j \) enlarges the dimension by \( (T_j - T) \) times. The UTLBP’s informativeness and distinctiveness is compared to the conventional LBP in TABLE 1. Assume that the LBP features of “original images” shown in the 2nd column are to be extracted. The specific portion of the image is magnified in the 3rd column and the center pixel grid mask is drawn on the image to simplify LBP feature extraction for human. Magnified images are shown in higher resolution to
be visible and the down-sampled images were used for feature extraction. Images #1 and #2 show samples of an African woman and a Caucasian man with mustache and beard, respectively. Although to human perception they are totally different the LBP feature extracted from the lips’ corner are thoroughly the same. Images #3 and #4, show light blue and dark brown samples among variety of eye colors [10] whose conventional LBP features are the same. Features from the grid center point of 4 samples in TABLE I using UTLBP are extracted in the last column with uniformly sampled thresholds (-30<Thr<30) and the step of 15. The UTLBP feature vector is represented by features of each threshold in a row. The UTLBP feature differentiates between image #1 and #2, the two totally dissimilar human faces and eye colors in image #3 and #4.

B. Uniform-sampling and Threshold Selection

Utilizing multiple thresholds to improve the encoding algorithm embedded in the LBP algorithm leads to larger feature dimension. However, the order of feature dimension increase is considerably smaller than increasing the radius (R) [8]. Like Gabor wavelet parameter adjustment [11], we need to select thresholds from the pool of selected uniformly-sampled thresholds in UTLBP to reduce the computational load. We suggest reducing the UTLBP feature dimension by analyzing the reliability of the feature set produced by each individual threshold. Then, choose highly reliable thresholds or range of reliable thresholds from the training set. We could reduce dimension furthermore by exploring the information diversity provided by each threshold, as proposed in [12] for Gabor wavelets.

<table>
<thead>
<tr>
<th>#</th>
<th>Original image</th>
<th>Magnified portion</th>
<th>Conventional LBP feature</th>
<th>Proposed UTLBP</th>
</tr>
</thead>
</table>
| 1  |                |                  | 01011000                  | [11111110 11111100 01011000 00001000 00000000] |}
| 2  |                |                  | 01011000                  | [11111110 11111110 01011000 00101000 01001000] |}
| 3  |                |                  | 11110111                  | [11110111 11110111 11110111 11000001 11000000] |}
| 4  |                |                  | 11110111                  | [11111111 11111111 11111111 11110111 11100000] |}

Figure 1. A facial image divided into 5×5 regions (b) The weights set for the weighted $\chi^2$ dissimilarity measure. Black squares indicate weight 0.0, dark gray 1.0, light gray 2.0 and white 4.0.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The LBP feature is well-known to be robust against illumination changes. Hence, we selected datasets with large variations to compare the performance of different versions of LBP in the literature and proposed in this paper. We compare them using face recognition and family verification datasets.

A. Face recognition

Among face recognition databases [13], CAS-PEAL provides large-scale face images with different sources of variations [14]. It contains 30,900 images of 1,040 subjects. Among them, 379 subjects have images with 6 different expressions. 438 subjects have images wearing 6 different accessories. 233 subjects have images between 10 to 31 lighting changes. 297 subjects have images against 2 to 4 different backgrounds colors. Furthermore, 66 subjects have images recorded in two sessions at a 6-month interval to provide aging effect. It also includes 21 poses for each subject. To prepare samples, faces are aligned and cropped to size 80×95 using provided eyes positions.

As proposed by Ahonen et al. in [8], faces are divided into regions (we selected 25 regions) as shown in Figure 1.a. LBP histograms of regions are extracted and using the proposed method in [15], the weight ($w_j$) for each region is calculated as shown in Figure 1.b. The weighted Chi square $\chi^2$ dissimilarity measure is employed as the distance measurement of the nearest neighbor classifier as in equation (7) [8],

$$\chi^2_w(x, \xi) = \sum_{j=1}^{J} w_j \frac{(x_{ij} - \xi_{ij})^2}{x_{ij} + \xi_{ij}}$$

where ‘x’ and ‘$\xi$’ are the normalized histograms to be compared and indices ‘i’ and ‘j’ refer to the $i$-th bin in histogram corresponding to the $j$-th local region ‘j’ with the weight ‘$w_j$’ calculated as for that region.

The CAS-Peal database contains Gallery (1040 images) and Training sets (1200 images) which are used as the training samples [14]. The Training set contains four images randomly selected from the frontal subset of 300 random subjects. Each sample in the probe set is tested against all
training samples to find the nearest neighbor. If the individual is recognized incorrectly it is considered as an error. We choose aging, lighting, background, accessory and distance probe sets of the CAS-Peal database as our work does not include pose correction and expression recognition. Among illumination normalization methods suggested in [14] as the standard procedure, histogram equalization is selected for image preprocessing. The histogram of 25 regions are calculated using LBP(8,1), LBP(12,2), LTP(8,1), 2nd order LDP and the proposed feature operator UTLBP(8,1) and UTLBP(12,2). To adjust thresholds, 200 random individuals of the FERET dataset [17] were tested for face recognition. The UTLBP features extracted by each threshold in the interval (-254<Thr<254) with the step equal to 5 were employed separately to perform face recognition for each threshold reliability analysis. The range of highly reliable thresholds were chosen that are (-69<Thr<69) with equal step of “5” for UTLBP in TABLE II and TABLE III. The weighted Chi square $\chi^2$ distance of UTLBP is measured for each threshold histogram subset and summed together. The recognition error rates of feature operators for selected probe sets are tabulated in Table II. We selected the first 1000 samples and the first 500 of the lighting and accessory probe sets respectively. All samples of other given probe sets were selected for evaluation.

As shown in TABLE II, UTLBP(12,2) outperforms other selected LBP features in the literature. Comparing to LBP(12,2) the significant amount of improvement (8%) occurs in the aging probe set which is one of the common changes in family. Other variations such as accessory, lighting, distance and background almost exist in family photo albums in which UTLBP achieves higher accuracy. This encourages us to benefit from UTLBP in family verification.

\[ \text{Recognition error rate for feature operators in each probe set of the CAS-PEAL database} \]

<table>
<thead>
<tr>
<th>Probe sets</th>
<th>Feature types recognition error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LBP(8,1)</td>
</tr>
<tr>
<td>Aging</td>
<td>0.37</td>
</tr>
<tr>
<td>Lighting</td>
<td>0.90</td>
</tr>
<tr>
<td>Background</td>
<td>0.034</td>
</tr>
<tr>
<td>Accessory</td>
<td>0.06</td>
</tr>
<tr>
<td>Distance</td>
<td>0.051</td>
</tr>
</tbody>
</table>

B. Family verification

Family verification task is to recognize if the query face belongs to the particular family [5]. Currently, to our knowledge, there is no publicly available dataset suitable for family verification except the single dataset published online [16]. Family photo albums searched from the web are the only available datasets to evaluate family verification. However, such family albums have various illumination conditions, backgrounds, accessories worn and the period of time elapsed between shots.

We collected 25 family datasets. Each dataset includes at least 10 samples per member (there are more samples in most cases) and totally contains about 90 samples on average. The overall dataset contains more than 2300 face images. The datasets cover various ethnicities, mixed ethnic families, with adult or very young children, parents with high aging effect and twin kids, shots taken at different years showing aging effect, with acceptable pose variations, in different locations, distances and lighting conditions. Aside from family samples 500 samples of the FERET dataset [17] and 500 faces of digital albums taken in the same conditions of family datasets from kids to adults with various ethnicities are collected as non-family samples. Faces are detected, aligned and cropped based on eyes coordinates and resized to size 80×95. The following scenarios are conducted to cover real applications:

Test 1: All family members are present in both training and test sets.

Test 2: A member is not present or known i.e. missing child or unknown father. Two different scenarios are defined: Test 2.a, the child with maximum number of samples is omitted from the training set. Test 2.b, one of the parents with maximum number of samples is omitted from the training set.

To perform family verification for family ‘i’, equal percentage (67%) of every member’s samples are selected as the ‘i-th’ family samples for training. All family faces are labeled as class “1” and the same ratio of the total number of non-family members are labeled as class “0”. The remaining samples were used as testing data. In test 2, 67% of all family members’ samples, except the omitted member, are used for training plus 67% of non-family members. The test set contains only the omitted member samples together with the remaining non-family members. Among feature listed in TABLE II, LDP, LBP and UTLBP are employed as they achieved higher accuracy compared to the rest in the face recognition test. The same algorithm as in [8], measuring the weighted Chi square $\chi^2$ distance for the nearest neighbor classification is utilized.

The family verification performance is measured by the False Negative Ratio, FNR, given by the number of family members incorrectly classified as non-family members over total number of family samples in the dataset, and False Positive Ratio, FPR, given by the number of non-family members incorrectly classified as family members over total number of non-family samples. The overall error or total error (Err) is calculated by total number of samples classified incorrectly over the total number of samples.
The family verification error is evaluated in TABLE III for selected features for test 1 and 2. The results show that UTLBP(12,2) reduces the total error (Err) by almost 1% comparing to LBP(12,2) in test 1. The 2nd order LBP was reported to outperform LBP(8,1) in [9] which is mostly consistent with TABLE II and TABLE III. The interesting point is, we expect to achieve higher accuracy by increasing the LBP operator radius ‘R’ and consequently number of sampling pixels ‘P’. In test 2.a and 2.b where we encounter a more challenging situation when a member is missing enlarging the radius from “1” to “2” leads to higher total error (Err) by 2% in test 2.a. Among selected features in the literature, LBP(12,2) achieved the least error rate in face recognition as stated in Table. II. Comparing to LBP(12,2), the proposed feature UTLBP(8,1) could reduce the family verification total error (Err) in case of the missing member by 3% and 2% in test 2.a and 2.b, respectively. Note that the same thresholds of UTLBP in Table II. were employed for feature extraction of family datasets. We employ the Genetic algorithm with Real-valued chromosome structure and diversity measurement cost function [12] to search for highly diverse thresholds on the training set. In this way, we can get the same accuracy of all thresholds by using only 7 to 10 thresholds depending on the family album. Hence, a reasonable trade off is obtained between the feature dimension and performance by the proposed UTLBP operator.

V. CONCLUSIONS

In this paper, we reviewed the LBP operator and its enhancements such as LBP(P,R), LTP and LDP. We highlighted that the relative pixel intensity quantization error in the LBP operator step function could lead to equivalent extracted feature despite of large images sparseness. We embed the missing information in the proposed Uniformly-sampled Thresholds for LBP (UTLBP) operator to augment feature distinctiveness. To examine the performance of features in the literature and the proposed UTLBP operator, various tests were conducted on the large face recognition CAS-PEAL database probe sets. The proposed operator outperforms other selected features in the literature up to 8% depending on the selected probe set. The proposed UTLBP(8,1) also improves the family verification performance up to 3% in the collected dataset when a member is missing where contrary to expectation the performance drops up to 2% by enlarging the LBP operator radius.

<table>
<thead>
<tr>
<th>Feature type</th>
<th>Error types for Test 1</th>
<th>Error types for Test 2.a</th>
<th>Error types for Test 2.b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Err</td>
<td>FPR</td>
<td>FNR</td>
</tr>
<tr>
<td>LDP (2nd order)</td>
<td>0.0500</td>
<td>0.0321</td>
<td>0.2067</td>
</tr>
<tr>
<td>LBP (8,1)</td>
<td>0.0518</td>
<td>0.0406</td>
<td>0.1470</td>
</tr>
<tr>
<td>LBP (12,2)</td>
<td>0.0483</td>
<td>0.0403</td>
<td>0.1281</td>
</tr>
<tr>
<td>UTLBP (8,1)</td>
<td>0.0506</td>
<td>0.0338</td>
<td>0.1471</td>
</tr>
<tr>
<td>UTLBP (12,2)</td>
<td>0.0391</td>
<td>0.0316</td>
<td>0.1312</td>
</tr>
</tbody>
</table>

REFERENCES


