RELAXED LOCAL TERNARY PATTERN FOR FACE RECOGNITION

Jianfeng Ren

BeingThere Centre
Institute of Media Innovation
Nanyang Technological University
50 Nanyang Drive, Singapore 637553.

Xudong Jiang, Junsong Yuan

School of Electrical & Electronics Engineering
Nanyang Technological University
50 Nanyang Avenue, Singapore 639798

ABSTRACT

Local binary pattern (LBP) is sensitive to noise. Local ternary pattern (LTP) partially solves this problem by encoding the small pixel difference into a third state. The small pixel difference may be easily overwhelmed by noise. Thus, it is difficult to precisely determine its sign and magnitude. In this paper, we propose the concept of *uncertain* state to encode the small pixel difference. We do not care its sign and magnitude, and encode it as both 0 and 1 with equal probability. The proposed Relaxed LTP is tested on the CMU-PIE database, the extended Yale B database and the O2FN mobile face database. Superior performance is demonstrated compared with LBP and LTP.

***Index Terms*—** Local Binary Pattern, Local Ternary Pattern, Uncertain State, Relaxed LTP, Face Recognition.

1. INTRODUCTION

Local binary pattern has been widely used in various computer vision applications because of its simplicity and robustness to illumination variations. However, its sensitivity to noise limits the performance. For local ternary pattern [1], the pixel difference between the center pixel and the neighboring pixel was encoded into a trinary code. LTP is less sensitive to noise as the small pixel difference is encoded into a separate state. To reduce the dimensionality, the ternary code is split into two binary codes: a positive LBP and a negative LBP.

Other than the ternary code, Nanni et al. proposed a quinary code and then split the quinary code into four binary codes [2]. In Local Adaptive Ternary Patterns [3] and extended LTP [4], instead of using a constant threshold, a threshold is calculated for each local window using local statistics, which makes LATP and ELTP less sensitive to noise and illumination variations. In Local Triplet Pattern [5], the equality is modeled as a separate state.

LTP partially solves the noise-sensitive problem by encoding the small pixel difference into a separate state. However, when the ternary code is split into a positive LBP code and a negative LBP code, it may result in a significant information loss. Furthermore, the positive and negative LBP histograms are strongly correlated, and hence a lot of redundant information may reside in those two histograms.

In this paper, the concept of *uncertain* state is introduced. If the pixel difference is within a threshold, it is encoded into the *uncertain* state. In this state, the sign and magnitude of the pixel difference cannot be precisely determined because it may be easily overwhelmed by noise. Based on the concept of *uncertain* state, we propose Relaxed LTP (RLTP). A ternary code is derived by encoding the large pixel difference into two strong states and encoding the small pixel difference into a separate *uncertain* state. Then the pixel difference that belongs to *uncertain* state is equally split into two strong states to represent the fact that the small pixel difference is equally likely to be positive and negative. In such a way, the ternary code is transformed back to a binary code. Different from LTP that results in two LBP histograms, the proposed RLTP results in only one LBP histogram. One ternary code with *uncertain* state contributes to more than one bin in the LBP histogram.

Many attempts were made to solve the face recognition problem, such as the linear [6, 7] and nonlinear [8] subspace approaches, and the local structure [9] and texture [1] based techniques. The proposed algorithm is firstly tested on CMU-PIE database with injected noise, and demonstrates strong resistance to noise compared with LBP and LTP. We further compare the proposed RLTP with LBP and LTP on other two challenging face databases: the extended Yale B database and the O2FN mobile face database. The proposed approach consistently outperforms LBP and LTP.

2. PROPOSED RELAXED LTP

Let us firstly briefly analyze the problem of LBP and LTP. Local binary pattern encodes the pixel differences between the neighboring pixels and the center pixel, e.g. $LBP_{P,R}$ encodes the pixel differences between the center pixel and P neighbors on a circle of radius R. Denote i_c as the gray level of the

This research is supported in part by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office.

center pixel, and i_p as the gray level of the circular neighbors, where $p = 0, 1, \ldots, P-1$. Then, the LBP code is obtained as: $LBP_{P,R} = \sum_{p=0}^{P-1} s(i_p - i_c)2^p$, where $s(z)$ is the threshold function: $s(z) = 1$ if $z \geq 0$, $s(z) = 0$ otherwise. LBP is sensitive to noise. A small image noise will cause the pixel difference encoded from 0 to 1 or vice versa.

LTP is less sensitive to noise, as it encodes the small pixel difference into a separate state. The LTP code is obtained as: $LTP_{P,R} = \sum_{p=0}^{P-1} s'(i_p - i_c)3^p$, where $s'(z,t)$ is the threshold function, and t is a pre-defined threshold.

$$s'(z,t) = \begin{cases} 1 & \text{if } z \geq t, \\ 0 & \text{if } |z| < t, \\ -1 & \text{if } z \leq -t. \end{cases} \tag{1}$$

The dimensionality of LTP histogram is very large. $LTP_{8,2}$ will result in a histogram of $3^8 = 6561$ bins. Thus, in [1], LTP code is split into a positive LBP code and a negative code as:

$$s'_p(z,t) = \begin{cases} 1 & \text{if } z \geq t, \\ 0 & \text{if } z < t. \end{cases} \tag{2}$$

$$s'_n(z,t) = \begin{cases} 1 & \text{if } z \leq -t, \\ 0 & \text{if } z > -t. \end{cases} \tag{3}$$

However, a significant amount of information may be lost during this process. On the other hand, the positive LBP histogram and the negative LBP histogram are strongly correlated, and hence a large mount of redundant information may reside in those two histograms.

In order to solve the problem of LBP and LTP, we propose Relaxed LTP. Firstly, let us redefine the thresholding function of local ternary pattern as:

$$s''(z,t) = \begin{cases} 1 & \text{if } z \geq t, \\ X & \text{if } |z| < t, \\ 0 & \text{if } z \leq -t. \end{cases} \tag{4}$$

State *0* and *1* represent two strong states where the pixel difference is almost definitely negative and positive, respectively. Noise can unlikely change it from 0 to 1 or from 1 to 0. State *X* represents an *uncertain* state where the pixel difference is small. The small pixel difference is vulnerable to noise if we only take its sign. More specifically, noise can easily change its LBP bit from 0 to 1 or vice versa. Therefore, regardless its sign and magnitude, we encode the small pixel difference into the *uncertain* state. Such a coding scheme is less sensitive to noise.

Tri-state code results in a histogram of too many bins, and hence it is desirable to reduce its dimensionality. Two strong states are corresponding to large pixel differences, which are less affected by image noise and more reliable. The small image difference in State *X* is easily distorted by noise, and hence less reliable. Thus, we aim to deduce State *X* into the strong states. It is difficult to precisely determine the sign and magnitude of small pixel difference. Based on our initial observations, we find that the small pixel difference is equally likely to be positive and negative. Thus, we encode State *X* equally into two strong states, i.e. State *0* and *1* with equal probability. As a result, we transform the trinary code back to the binary code.

The encoding process is shown in Fig. 1. The image patch is firstly encoded as a trinary code. Then, the trinary code is encoded into binary codes, e.g. bit 0 and bit 5 are encoded as both *0* and *1*. In this example, the ternary code results in four LBP codes. Those four binary codes are equally likely. Therefore, when constructing the histogram, each binary code contributes 0.25 to the corresponding bin of the histogram.

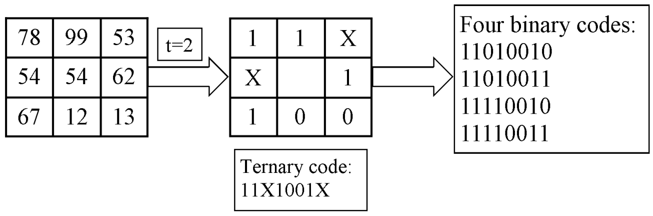


Fig. 1. Illustration of encoding scheme of Relaxed LTP. Firstly, the trinary code *11X1001X* is derived from the image. When constructing the histogram, it results in four binary codes.

Now we describe the general form of the proposed Relaxed LTP encoding scheme. Instead of hard code the pixel difference z as *0* or *1*, we assign the probability of encoding it as *0* or *1*. Given the pixel difference z, denote $P^1(z)$ as the probability to encode it as 1, and $P^0(z) = 1 - P^1(z)$ as the probability to encode it as 0. For the proposed Relaxed LTP, the *uncertain* state is encoded as both *1* and *0* with equal probability. Therefore, Eqn. 4 can be reformulated as:

$$P^1(z) = \begin{cases} 1 & \text{if } z \geq t, \\ 0.5 & \text{if } |z| < t, \\ 0 & \text{if } z \leq -t. \end{cases} \tag{5}$$

When constructing the resulting LBP histogram, we calculate the probabilities of all 256 patterns as:

$$P_j = \prod_{i=0}^{7} c_i P_i^1(z) + (1 - c_i)P_i^0(z), \tag{6}$$

where the LBP code $j = \sum_{i=0}^{7} c_i * 2^i$; c_i is *i-th* bit of the code; $P_i^1(z)$ and $P_i^0(z)$ are the probabilities that bit *i* should be encoded as *1* and *0*, respectively. In practice, for every trinary pattern, the probabilities of all 256 patterns can be pre-calculated and stored in a look-up table. The probabilities of all the pixels within one patch are summed up to form the LBP histogram of the patch.

The proposed RLTP is significantly different from LTP. When there is only one *uncertain* bit, both the proposed RLTP and LTP produce two LBP codes. However, those two codes are accumulated in two different histograms for LTP, whereas in one histogram for RLTP. When there are more *uncertain* bits, the proposed RLTP differs more significantly from LTP. Take ternary code *11X1001X* in Fig. 1 as an example. For LTP it results in two LBP codes: *11010010* and *00001100* (its complement is *11110011*), which are accumulated in the positive LBP histogram and the negative LBP histogram, respectively. For the proposed approach, four binary codes are produced as shown in Fig. 1, in which the first and the fourth code are corresponding to the positive LBP code and the complement of the negative LBP code of LTP, respectively. Those four codes are accumulated in one LBP histogram. Each code contributes 0.25 to the corresponding bin of the histogram.

3. EXPERIMENTAL RESULTS

In this section, we compare the proposed approach with LBP and LTP on the CMU-PIE database [10], the extended Yale B face database [11, 12] and the O2FN mobile face database [7]. All the images are geometrically normalized so that the eyes are at the same position, similarly as in [1]. The normalized image is of size 128×128 pixels. Each image is divided into patches of size 8×8 pixels. $LBP_{8,2}$, $LTP_{8,2}$ and $RLTP_{8,2}$ are used. The histogram is calculated for each patch. Each RLTP histogram has 256 bins. Chi-square distance is utilized to calculate the dissimilarity score and the nearest neighbor classifier is used. In order to reduce illumination variations, the images are pre-processed similarly as in [1], i.e. Gamma correction, Difference of Gaussian filtering and contrast equalization. We utilize the source codes provided by the authors of [1] to perform photometric normalization.

3.1. Experimental Results on the CMU-PIE Database

The CMU-PIE database contains over 40000 facial images of 68 subjects, across 13 different poses, under 43 different illumination conditions, and with 4 different expressions. We choose the illumination set for our experiments, which contains 1407 images of 67 subjects.[1] Each subject has 21 images. We adopt a challenge experimental setting, in which only the image with frontal lighting (Image ID 08) is used as the gallery set and the rest images with large illumination variations are used as the probe set. We compare the proposed approach with LBP and LTP on images injected with noise. The images are normalized in the range of $(0, 1)$, and then we apply additive uniform noise in the range of $(-p/2, p/2)$. The corresponding standard derivation can be derived as $\sigma_u = p/\sqrt{12}$. We vary the noise range for $p = 0.1, 0.15, 0.2$, and respectively $\sigma_u = 0.0289, 0.0433, 0.0577$. The sample images are shown as the first row of Fig. 2 and the photometrically normalized images are shown in the second row.

The recognition rates of LBP, LTP and RLTP are shown in Fig. 3. For different noise settings and different thresholds, the proposed Relaxed LTP consistently outperforms LBP and LTP. The highest recognition rates and the optimal thresholds under different noise settings for LBP, LTP and RLTP are summarized in Table. 1. When the noise level increases, the recognition rates of LBP and LTP drop significantly, whereas RLTP still preserves a high recognition rate. The proposed approach demonstrates strong resistance to noise compared with LBP and LTP. Furthermore, the optimal threshold for LTP varies with noise level, and hence it is difficult to determine an optimal threshold for LTP. In contrast, the optimal threshold for RLTP is consistent over different noise settings, i.e. $t = 2$.

[1]The images of Subject 39 are not complete and hence excluded.

Fig. 2. The first row shows the sample images of CMU-PIE with additive uniform noise $p = 0, 0.1, 0.15, 0.2$. The second row shows the photometrically normalized images.

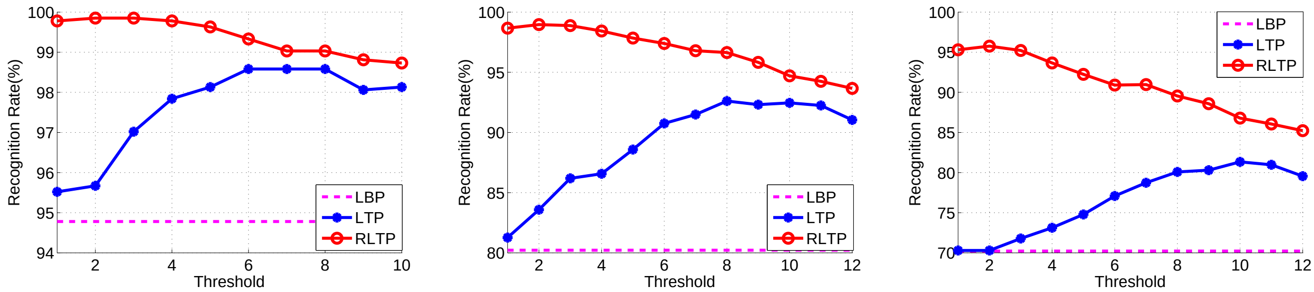
Method	$p = 0.1$	$p = 0.15$	$p = 0.2$
LBP	94.78%	80.22%	70.22%
LTP	98.58% (6)	92.61% (8)	81.34% (10)
RLTP	99.85% (2)	98.96% (2)	95.75% (2)

Table 1. The highest recognition rates and the optimal thresholds under different noise settings for LBP, LTP and RLTP on the CMU-PIE database.

3.2. Experimental Results on the Extended Yale B Database

The extended Yale B database [11, 12] contains 38 subjects under 9 poses and 64 illumination conditions. We follow the same database partition as in [1], the images with most neutral light sources("A+000E+00") are used as the gallery images and all other frontal images are used as the probe images (In total 2414 images of 38 subjects). The database contains large illumination variations. Some images are taken under extreme lighting conditions.

The recognition rates for LBP, LTP and Relaxed LTP under different thresholds are shown in Fig. 4. RLTP consistently outperforms LBP and LTP. The highest recognition rates and the optimal thresholds are summarized in Table. 2. Compared with LBP and LTP, RLTP increases the recognition rate from 96.07% and 98.25% to 98.71%, respectively.



(a) $p = 0.1$ (b) $p = 0.15$ (c) $p = 0.2$

Fig. 3. Recognition rate of LTP and Relaxed LTP v.s different thresholds for different noise settings. The recognition rate of LBP is plotted in dotted line. RLTP consistently outperforms LBP and LTP.

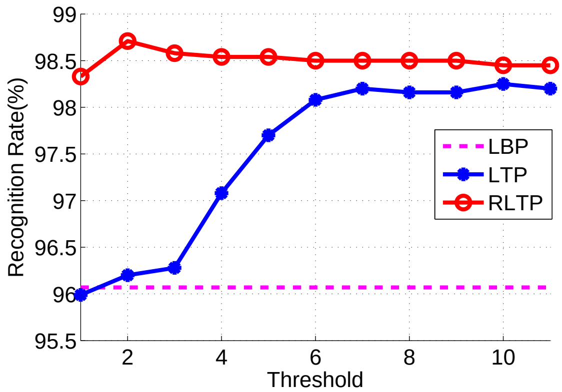


Fig. 4. Recognition rates of LBP, LTP and RLTP for different thresholds on the extended Yale B database.

Method	LBP	LTP	RLBP
Recognition rate	96.07%	98.25% (10)	**98.71%** (2)

Table 2. The highest recognition rates and the optimal thresholds for LBP, LTP and RLTP on the Yale B database.

3.3. Experimental Results on the O2FN Mobile Database

The O2FN mobile face database [7] is designed to evaluate the face recognition algorithms on images captured by mobile phone, which are of low resolution and low image quality. It contains 2000 face images of size 144×176 pixels for 50 subjects. The database mainly contains in-plane rotations and illumination variations. The first 20 images of each subject are chosen as the gallery images and the rest are used as the probe images.

The recognition rates for LBP, LTP and RLTP under different thresholds are shown in Fig. 5. RLTP consistently outperforms LBP and LTP. The highest recognition rates and the optimal thresholds are summarized in Table. 3. Compared with LBP and LTP, RLTP increases the recognition rate from 96.60% and 97.40% to 98.40%, respectively.

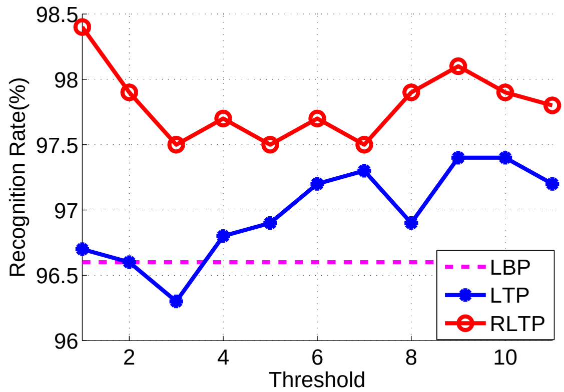


Fig. 5. Recognition rates of LBP, LTP and RLTP for different thresholds on the O2FN database.

Method	LBP	LTP	RLBP
Recognition rate	96.60%	97.40% (9)	**98.40%** (1)

Table 3. The highest recognition rates and the optimal thresholds for LBP, LTP and RLTP on the O2FN database.

4. CONCLUSION

In this paper, we address the challenge of improving the robustness to image noise. We propose the concept of *uncertain* state, in which the small pixel difference is encoded as both 0 and 1 with equal probability. The proposed Relaxed LTP shows strong resistance to noise compared with LBP and LTP. On the CMU-PIE database, the extended Yale B database and the O2FN database, the proposed RLTP consistently outperforms LBP and LTP. Furthermore, the optimal threshold for RLTP is more consistent compared with LTP. The optimal thresholds for LTP are ranging from 6 to 10 for different databases. In contrast, RLTP achieves the highest recognition rate consistently at a low threshold. In general, the threshold of RLTP can be selected as 2.

5. REFERENCES

[1] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635–1650, 2010.

[2] L. Nanni, A. Lumini, and S. Brahnam, "Local binary patterns variants as texture descriptors for medical image analysis," *Artificial intelligence in medicine*, vol. 49, no. 2, pp. 117–125, 2010.

[3] M.A. Akhloufi and A. Bendada, "Locally adaptive texture features for multispectral face recognition," in *IEEE International Conference on Systems Man and Cybernetics (SMC)*. IEEE, 2010, pp. 3308–3314.

[4] W.H. Liao and T.J. Young, "Texture classification using uniform extended local ternary patterns," in *2010 IEEE International Symposium on Multimedia (ISM)*. IEEE, 2010, pp. 191–195.

[5] D. He and N. Cercone, "Local triplet pattern for content-based image retrieval," *Image Analysis and Recognition*, pp. 229–238, 2009.

[6] X. Jiang, B. Mandal, and A. Kot, "Eigenfeature regularization and extraction in face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 3, pp. 383–394, 2008.

[7] Jianfeng Ren, Xudong Jiang, and Junsong Yuan, "A complete and fully automated face verification system on mobile devices," *Pattern Recognition*, vol. 46, no. 1, pp. 45 – 56, 2013.

[8] Xudong Jiang, Bappaditya Mandal, and Alex Kot, "Complete discriminant evaluation and feature extraction in kernel space for face recognition," *Machine Vision and Applications*, vol. 20, no. 1, pp. 35–46, 2009.

[9] Cong Geng and Xudong Jiang, "Face recognition based on the multi-scale local image structures," *Pattern Recognition*, vol. 44, no. 10, pp. 2565–2575, 2011.

[10] T. Sim, S. Baker, and M. Bsat, "The cmu pose, illumination, and expression (pie) database," in *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*. IEEE, 2002, pp. 46–51.

[11] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, 2001.

[12] K.C. Lee, J. Ho, and D.J. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 684–698, 2005.