Noise-resistant local binary pattern based on random projection

Shasha Li
lishasha6464@163.com

Yukai Tu
tyk@bupt.edu.cn

Weihong Deng
whdeng@bupt.edu.cn

Beijing University of Posts and Telecommunications
No.10 Xitucheng Road, Haidian District, Beijing, PR China

Jiwen Lu
Jiwen.Lu@adsc.com.sg
Tsinghua University
Tsinghua Campus, Haidian District, Beijing, PR China

Abstract

Local binary pattern (LBP) is sensitive to noise. LBP projects local patch to eight-dimension vector by operating subtractions between pixel and its neighborhood. Two adjacent pixel values are generally very close, thus little noise can change their relative magnitude, leading to coding err. Using another projecting approach, random projection, as an alternate, we propose local binary pattern based on random projection (RPLBP), which is much more robust to noise while remaining computationally simple. Experiments on the aligned FERET database show that RPLBP has outstanding robustness to variations in illumination, facing expressions and aging. It's worth mentioning that the proposed RPBLP demonstrates superior noise-resistant performance to LBP and NRLBP.

1. Introduction

Face recognition has become an important topic in computer vision with more than 30 year's development. Among local face descriptors, LBP [2], [3] has emerged as one of the most prominent and has been widely used in texture classification [14], facial analysis [2] and many other tasks [15].

LBP encodes the signs of the pixel differences between a pixel and its neighboring pixels to a binary code [16]. From other points of view, this encoding procedure is also a projecting procedure. LBP uses projection matrix ϕ to project patch vector P to eight dimensional vector X by operating subtractions between pixel and its neighborhood. Employ binarization to X then we can get LBP label. LBP has outstanding advantages. However, LBP has prominent limitations: 1.sensitive to noise; 2.failing to capture large-scale information[14]. The sensitiveness to noise is essentially for the reason that the projection contains subtractions between two adjacent pixel values, which are usually very close and little noise can change their relative magnitude, leading to coding err. Actually, the ϕ in LBP approach is a bad projection matrix, which is closely associated with the two drawbacks of LBP.

Random projection (RP) [13], [9], [8] refers to the technique of projecting a set of points from a high dimensional space to a randomly chosen low-dimensional subspace. The information-preserving and dimensionality-reduction power of RP is firmly evidenced in the emerging theory of compressed sensing (CS) [6], [7], [12], which states that, for sparse and compressible signals, a small number of non-adaptive linear measurements in the form of random projections can capture most of the salient information in the signal and allow for perfect reconstruction of the signal.

Random projection is a wonderful alternate to the bad projection in traditional LBP method, so in this paper we propose a simple, efficient, yet robust local descriptor for face recognition, local binary pattern based on random projection (RPLBP). The proposed descriptor is information-preserved, very compact while remaining robust to noise.

2. Local binary pattern based on random projection (RPLBP)

2.1. Random projection

Denote random production as follows.

$$X = \phi P \quad X \in R^{m\times 1}, P \in R^{n\times 1}, m \ll n \qquad (1)$$

Finding the information reserving $\phi \in R^{m\times n}$ which can map from the sparse to the measurement domain is the thing we should care about. Information preserving means that ϕ

provides a stable embedding that approximately preserves distances between all pairs of signals, that is, for any two patches, P_1 and P_2, the distance between them is approximately preserved. We express this formally in Eq.2.

$$1-\epsilon \leq \frac{\|\phi(P_1-P_2)\|}{\|P_1-P_2\|} \leq 1+\epsilon \tag{2}$$

Johnson-Lindenstrauss lemma [9] states the existence of information preserving ϕ. It points out that Eq.2 can be satisfied with overwhelming probability with those ϕ, whose entries are randomly drawn from certain probability distributions [1], specifically including the Gaussian distribution [1], [5], [4], which means random Gaussian projection approximately preserves pairwise distances in the data set, so this paper use Gaussian matrix as the random projecting matrix.

2.2. RPLBP operator

For one pixel location, a patch P of size $\sqrt{n} \times \sqrt{n}$ is extracted around it. Random projection is then used to set up projection to the patch to gain random measurement as Equ3. The entries of ϕ are sampled from independent zero-mean, unit-variance normal distribution.

$$X = \phi P, \phi \in R^{8\times n} \tag{3}$$

RPLBP can then be gained by thresholding 8 elements of X with zero and considering the result as a binary number, which can be donated as follows.

$$C = \{c = \frac{1+sgn(x)}{2} | x \in X\} \tag{4}$$

where

$$sgn(x) = \begin{cases} 1 & x \geq 0, \\ -1 & x < 0. \end{cases}$$

The 8-digit binary number of each pixel is the RPLBP symbol, which can also be transformed to decimal number between 0 and 255.

2.3. Proposed Approach

1. **Description of face image based on RPLBP**

   Step1. A symbol is assigned to every pixel of an image.

   The image is processed by extracting patches of size $\sqrt{n}\times\sqrt{n}$ around each pixel position. By using RPLBP operator on each patch, we then transform the image into image of RPLBP symbols.

   Step2. Regional descriptors are constructed

   Facial image is divided into local regions R_1, R_2,$\cdots$ Symbols over each region are then summed up to compute a histogram for regional description.

   Step3. Global descriptor is constructed

   All the regional histograms are concatenated to form global histogram, which encodes both the appearance and the spatial relations of facial regions.

2. **Classifier**

   We consider two kinds of classifier.

   - Nearest Neighbor Classifier (NNC). The distance between two histograms is measured using the χ^2 statistic.

   $$\chi^2(h_1-h_2) = \frac{1}{2}\sum_{k=1}^{K}\frac{(h_1(k)-h_2(k))^2}{h_1(k)+h_2(k)} \tag{5}$$

   - Linear Regression Analysis (LRA) [11]. For the recognition of C subjects, we first have to study an optimal transformation matrix W based on the gallery set.

   $$W = YX^H \tag{6}$$

   where Y is a $C \times C$ unit matirx, X is the gallery samples $X = [x_1; x_2; \cdots x_C]$, X^H denotes the generalized inverse of X.

   When a novel test image is presented to the LRA classifier, the feature vector of the image donated by x_{new} is first extracted and then normalized to zeros mean and unit length. The response vector y is derived by a linear transformation $y = Wx$. Finally, the recognition result is determined by the largest component of the response vector.

   $$w = \arg\max_{i=1,\ldots,C} y^{(i)} \tag{7}$$

3. Symbol err rate of RPLBP and LBP under addictive Gaussian noise

Gaussian noise is one of the most common kinds of noise, so this part compares the anti-Gaussian-noise ability of RPLBP with LBP. The patch size of RPLBP method is 3×3 pixels, which is the same to LBP method.

The only difference between RPLBP and LBP method is ϕ in Equ.3. Rewrite Equ.3 as follows.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_8 \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{19} \\ \alpha_{21} & & & \\ \vdots & & & \\ \alpha_{81} & \cdots & & \end{pmatrix} \cdot \begin{pmatrix} p_1+n_1 \\ p_2+n_2 \\ \vdots \\ p_9+n_9 \end{pmatrix} \tag{8}$$

where $(\alpha_{i1}, \alpha_{i2}, \alpha_{i3} \cdots \alpha_{i,9}), i = 1, 2 \cdots 8$ is normalized vector; $(n_1, n_2 \cdots n_9)$ is addictive Gaussian noise and $n \sim N(0, \sigma^2)$.

- For LBP, ϕ in Equ.8 is as follows.

$$\phi = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} & \cdots & 0 \\ \vdots & & & & \\ \frac{1}{\sqrt{2}} & 0 & 0 & \cdots & \frac{-1}{\sqrt{2}} \end{pmatrix} \tag{9}$$

- For RPLBP, entries of ϕ in Equ.8 are sampled from independent zero-mean, unit-variance normal distribution.

For every x, there are two part of it. Taking x_1 for example, x_{1r} is the exact x_1 when there is no noise; x_{1n} represents the impact of noise.

$$\begin{aligned} x_1 &= x_{1r} + x_{1n} \\ &= (\alpha_{11}p_1 + \alpha_{12}p_2 + \cdots + \alpha_{19}p_9) + \\ &\quad (\alpha_{11}n_1 + \alpha_{12}n_2 + \cdots + \alpha_{19}n_9) \end{aligned} \tag{10}$$

We can see from Equ.4 that the exact value of x doesn't matter. The code c will be right as long as the eight elements of x don't change from negative to positive or from positive to negative under noisy circumstances, that is to say, encoding wrong means under the impact of x_n, x_r changes its sign, thus if the probability density function of X_n is $n(x)$ and the probability density function of X_r is $r(x)$, calculate the bit err rate (BER) as Equ.11.

$$Err(\sigma) = \int_{-\infty}^{+\infty} n(x,\sigma) \int_{0}^{\|x\|} r(h)\mathrm{d}h\mathrm{d}x \tag{11}$$

3.1. The distribution of X_n

$$X_n = \alpha_1 N_1 + \alpha_2 N_2 + \cdots + \alpha_9 N_9 \tag{12}$$

N_i are independent random variables, $n \in N(0,\sigma^2)$ and $\|(\alpha_1, \alpha_2 \cdots \alpha_9)\| = 1$, thus X_n follows Gaussian distribution, $x_n \in N(0,\sigma^2)$.

3.2. The distribution of X_r

After statisticing 1000 face images, get the probability density function of X_r shown in Fig.1. Utilizing Equ.11 and Fig.1, get BER of RPLBP and LBP shown in Fig.2.

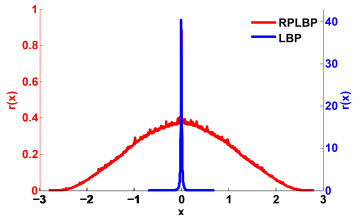


Figure 1: Probability density function of X_r.

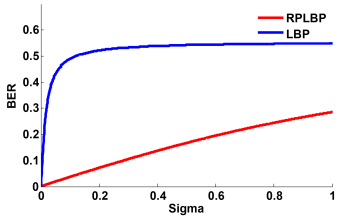
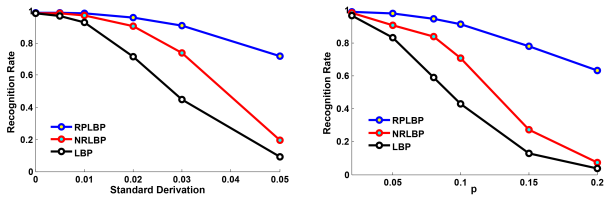


Figure 2: BER in theory



(a) RPLBP (b) LBP

Figure 3: SER of RPLBP and LBP respectively in theory and in practice

It is easy to see from Equ.11 that if the distribution of X_r is concentrated around zero value, the integral $\int_0^{\|x\|} r(h)\mathrm{d}h$ will be large and the bit err rate will be high. Thus the ϕ of LBP is a bad projection matrix, which makes result data too concentrated; in contrast, random projection is a good projection which is much more robust to noise.

3.3. Symbol err rate (SER) of RPLBP and LBP

$$SER = 1 - (1 - BER)^8 \tag{13}$$

We can calculate SER as in Equ.13. The SER of RPLBP and LBP in theory and in practice are shown in Fig.3(a) and Fig.3(b). SER of LBP method is almost 100 percent when σ reaches 0.04. However, for RPLBP, the SER doesn't reach 100 percent before σ reachs 1.

4. Experimental results

4.1. Experimental setup

Figure 4: (a) The images with additive Gaussian noise of $\sigma = 0.01, 0.02, 0.05$ respectively. (b) The images with uniform noise of $p = 0.05, 0.1, 0.2$ respectively.

This paper uses aligned FERET database[10] for experiment. There are 14,051 gray-scale images representing

1,190 individuals which contain variations in illustration, facing expressions, aging, etc. We used frontal faces in this experiment, which can be divided into five sets as follows.

- fa set (1,196 images), training set;
- fb set (1,195 images), probe set, people with alternative facial expressions than in the fa photograph;
- fc set (192 images), probe set, photos taken under different lighting conditions;
- dup I set (772 images), probe set, photos taken later in time;
- dup II set (234 images), probe set, photos taken at least a year after the corresponding gallery image.

To conduct experiments about capability of noise resistance, inject Gaussian and uniform noise of various noise levels on to the image of the probe set fa, fb, dup I and dup II. The simple of noisy images are shown in Fig.4

1. Adding Gaussian Noise: Normalize the images in range of $(0, 1)$, and then apply additive Gaussian noise with zero mean and standard derivation of σ.
2. Adding Uniform Noise: Uniform noise is another common type of noise. As the same with adding Gaussian noise, we conduct normalization and then apply additive uniform noise in the range of $(-p/2, p/2)$.

4.2. Parameters of the RPLBP method

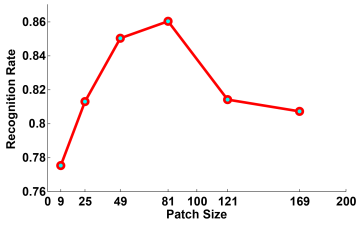
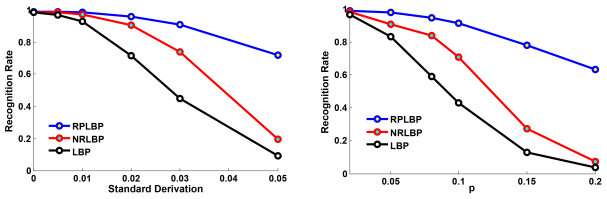


Figure 5: Recognition results on data set dup I as a function of patch size $\sqrt{n} \times \sqrt{n}$ for region size 15×13 and LRA classier

The size of patch and the division of the image can be chose to optimize the performance of the face recognition. Here, 15×13 pixel regions are used since it is a good trade-off between recognition performance and feature vector length. Fig.5. plots recognition rate over patch size n. When n is small, the performance improves rapidly with the n increasing because large patches can capture large-scale structures that may be the dominant feature, are more robust to variations. However, when n gets too large, performance gets worse because eight binaries are not enough to preserve the information of that large patches.



(a) Gaussian noise (b) Uniform noise

Figure 6: Recognition results of RPLBP, LBP and NRLBP with Gaussian noise and uniform noise respectively on data set fb.

4.3. Comparing RPLBP to LBP and NRLBP

Table 1: Recognition rates of RPLBP approach and LBP approach using LRA and NNC classier

		fb	fc	dup I	dup II
LRA	RPLBP	98.66%	97.42%	85.04%	81.20%
	LBP	98.58%	94.33%	81.99%	75.64%
NNC	RPLBP	93.97%	54.12%	66.34%	54.70%
	LBP	92.64%	44.33%	57.76%	37.61%

Tab.1. shows the recognition results on aligned FERET test sets. As is seen, RPLBP method performs better than LBP when there are variations in facing expressions, lighting conditions and aging. The parameters of RPLBP are patch size $n = 9 \times 9$ pixels and region size 15×13 pixels. The parameters of LBP are region size 15×13 pixels and LBP operator $LBP^{U_2}_{8,2}$.

Table 2: Recognition rates with Gaussian noise

		$\sigma = 0.01$	$\sigma = 0.03$	$\sigma = 0.05$
fb	RPLBP	98.44±0.17%	90.67±0.46%	71.45±0.67%
	NRLBP	97.17±0.46%	73.74±0.25%	19.58±0.76%
	LBP	92.73±0.13%	44.96±1.01%	9.27±0.67%
fc	RPLBP	96.08±0.51%	88.53±1.29%	56.29±3.35%
	NRLBP	91.85±1.03%	49.18±1.81%	7.42±1.81%
	LBP	86.39±2.06%	15.46±3.09%	0.93±0.26%
dup I	RPLBP	83.74±0.41%	66.59±0.49%	41.02±1.46%
	NRLBP	76.73±0.69%	34.49±0.63%	6.04±0.42%
	LBP	69.20±0.76%	20.25±0.21%	3.82±0.42%
dup II	RPLBP	78.89±1.28%	59.29±2.53%	28.80±1.71%
	NRLBP	66.84±2.14%	17.09±0.86%	4.45±0.83%
	LBP	60.68±0.64%	10.94±1.07%	3.59±0.43%

We add NRLBP[16] in to verity the prominent noise-resistant ability of RPLBP because NRLBP is probably the most noise-resistant derivative LBP. Recognition results are

Table 3: Recognition rates with uniform noise

		$p = 0.05$	$p = 0.10$	$p = 0.15$
fb	RPLBP	97.80±0.16%	91.28±0.54%	78.08±1.01%
	NRLBP	90.78±0.16%	70.66±0.50%	27.42±0.46%
	LBP	83.18±0.88%	42.93±0.38%	12.97±0.29%
fc	RPLBP	95.88±0.51%	90.41±1.03%	69.33±2.06%
	NRLBP	84.02±2.32%	45.77±1.03%	14.23±2.06%
	LBP	68.15±0.77%	15.08±1.03%	1.68±1.03%
dup I	RPLBP	81.86±0.56%	68.00±1.04%	48.82±0.76%
	NRLBP	64.79±1.25%	31.77±1.04%	8.65±0.56%
	LBP	54.68±1.11%	18.60±0.83%	4.78±0.35%
dup II	RPLBP	74.78±0.86%	64.02±1.93%	39.32±0.42%
	NRLBP	54.53±1.71%	16.67±1.29%	4.70±1.29%
	LBP	43.59±2.14%	10.69±0.64%	4.38±0.22%

showed in Tab.2. and Tab.3. We plot Fig.6. to compare them more clearly, from which we can see RPLBP yields much higher recognition rates.

5. Conclusion

In this paper we have presented RPLBP, a theoretically and computationally simple, noise tolerant approach for face recognition, which has the following outstanding advantages:

1. Information-preserving property and low computational complexity. which are the innate superiority of random projection.

2. Outstanding robustness to additive noise. Random projection chooses a better projection direction, thus the mapped data are not so concentrated. Good projection guarantees the excellent robustness.

3. Capability of capture large-scale-information. Random matrix can be arbitrary size, thus the extracted patch can be large enough to capture more information.

6. Acknowledgements

This work was partially sponsored by supported by the NSFC (National Natural Science Foundation of China) under Grant No. 61375031, No. 61573068, No. 61471048, and No.61273217, the Fundamental Research Funds for the Central Universities under Grant No. 2014ZD03-01, This work was also supported by the Beijing Higher Education Young Elite Teacher Program, and the Program for New Century Excellent Talents in University.

References

[1] D. Achlioptas. Database-friendly random projections. *Proceedings of the twentieth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 274–281, 2001. 2

[2] T. Ahonen, A. Hadid, and M. Pietik01inen. Face recognition with local binary patterns. *Lecture Notes in Computer Science*, pages 469–481, 2004. 1

[3] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 28(12):2037–2041, 2006. 1

[4] R. Baraniuk, M. Davenport, R. Devore, M. Wakin, R. Baraniuk, M. Davenport, R. Devore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, 2008. 2

[5] G. Biau, L. Devroye, and G. Lugosi. On the performance of clustering in hilbert spaces. *IEEE Transactions on Information Theory*, 54(2):781 – 790, 2008. 2

[6] E. Candes and T. Tao. Decoding by linear programming. *Information Theory IEEE Transactions on*, 51(12):4203–4215, 2005. 1

[7] E. J. Candes and T. Tao. 2006), near-optimal signal recovery from random projections: universal encoding strategies. *IEEE Trans.inform.theory*, 52(12):5406 – 5425, 2006. 1

[8] S. Dasgupta. Experiments with random projection. In *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*, pages 143–151. Morgan Kaufmann Publishers Inc., 2000. 1

[9] S. Dasgupta and A. Gupta. An elementary proof of a theorem of johnson and lindenstrauss. random struct. In *Algorithms*, 2003. 1, 2

[10] W. Deng, J. Hu, J. Lu, and J. Guo. Transform-invariant pca: A unified approach to fully automatic facealignment, representation, and recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 36(6):1275–1284, 2014. 3

[11] W. Deng, J. Hu, X. Zhou, and J. Guo. Equidistant prototypes embedding for single sample based face recognition with generic learning and incremental learning. *Pattern Recognition*, 47(12):3738–3749, 2014. 2

[12] D. L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, 2006. 1

[13] W. B. Johnson and J. Lindenstrauss. Extensions of lipschitz mappings into a hilbert space. *Contemporary mathematics*, 26(189-206):1, 1984. 1

[14] L. Liu, B. Yang, P. Fieguth, Z. Yang, and Y. Wei. Brint: A binary rotation invariant and noise tolerant texture descriptor. In *Image Processing (ICIP), 2013 20th IEEE International Conference on*, pages 255–259, 2013. 1

[15] L. Nanni. Alessandra lumini, shery brahnam,local binary patterns variants as texture descriptors for medical image analysis. *Artificial Intelligence in Medicine*, 49(2):117C125, 2010. 1

[16] J. Ren, X. Jiang, and J. Yuan. Noise-resistant local binary pattern with an embedded error-correction mechanism. *Image Processing IEEE Transactions on*, 22(10):4049 – 4060, 2013. 1, 4