This letter is published by IET, Electronics Letters. The copyright has been transferred to IET. The final letter DOI: https://doi.org/10.1049/el.2018.7631

Local ZigZag Max Histograms of Pooling Pattern for Texture Classification

Swalpa Kumar Roy, Shiv Ram Dubey and Bidyut B. Chaudhuri

The efficiency of any texture classification model confides on descriptor used for similarity matching. The formation of image descriptor is a challenging and important task in computer vision. This letter introduces a Local ZigZag Max Histograms of Pooling Pattern (LZMHPP) for classification of texture images. To compute the descriptor, first the dissimilarity between center pixel and it's neighbors is computed for each image patch over the whole image, then the dissimilarity map is encoded with different type of ZigZag ordering mechanism, and finally the Max Histograms Pooling is used to form LZMHPP descriptor from two complementary ZigZag weighted structures and achieves sufficient robustness under geometric variations. The experimental study on KTH-TIPS and CUReT texture databases indicates the efficiency and supremacy of LZMHPP descriptor for texture classification.

Introduction: The texture recognition plays a vital role in many applications such as industrial image analysis, remote sensing image analysis, facial analysis, etc. The texture classification from images is a very challenging problem due to the huge amount of inter-class similarities. The primary method to deal with this problem is to match the descriptors computed from the texture images instead of image itself, as the descriptors exhibit the robustness and discriminative features. Several methods were proposed in the past to design the suitable image descriptors for the texture characterisation. Local binary pattern (LBP) is one of state-of-the-art descriptor for texture classification [1]. Several improvements over LBP have been proposed such as Local ternary pattern (LTP) [2], Local derivative pattern (LDP) [3], Local tetra pattern (LTrP) [4], Local bit-plane dissimilarity pattern (LBDISP) [5], Complete dual-cross pattern (CDCP) [6], etc. Recently, ZigZag based descriptor has shown a great improvement in the performance using Local directional ZigZag pattern (LDZP) [7]. The existing descriptors suffer due to the limited descriminative power and robustness ability.

In this letter, a new descriptor is proposed by exploiting the benefits of ZigZag features as well as Max histogram pooling. Basically, first different local ZigZag patterns are extracted, then the Max pooling is applied over the histograms of different local ZigZag patterns. The utilisation of different ZigZag features boost the discriminative ability, whereas the Max pooling of histograms introduces more robustness against the various geometrical variations such as rotation, scaling, etc.

Local ZigZag Max Histograms of Pooling Pattern: The local binary pattern (LBP) [1] is a well-known and highly discriminative texture descriptor which labels each pixel within a local neighborhood at a radius around its center by computing the *sign* of difference of the intensity values of neighbours from the center. The LBP achieves enough discriminate under different geometric variations. However, in LBP structure, there is no perceptual angular relation between two consecutive neighbouring samples with respect to its center. So, it can not capture enough angular information in the local texture patch which could be useful to capture the complex local structure. To overcome the aforementioned problems, we have designed a new descriptor in this letter based on local ZigZag weighed structure [7].

In LZP descriptor, the *sign* of intensity differences of the center to its neighboring pixels are calculated to label each pixel initially as,

$$\begin{aligned}\mathbf{f} &= [f_0, \dots, f_n, \dots, f_{N-1}]^T \\ &= [I_0 - I_c, \dots, I_n - I_c, \dots I_{N-1} - I_c]\end{aligned} \tag{1}$$

where $\mho \in \mathbb{R}^{1\times N}$ represents a vector of local intensity difference, which produces LZP invariant under monotonic photometric changes, I_c, I_n and N represent the pixel intensity of center, n^{th} neighbouring and total number of involved points (i.e, $N = 8$) respectively. The value of LZP at I_c is calculated as,

$$\text{LZP}(i,j) = \sum_{n=0}^{N-1} h(f_n) \times 2^n, \quad h(z) = \begin{cases} 1, & if\ z \geq 0 \\ 0, & else. \end{cases} \tag{2}$$

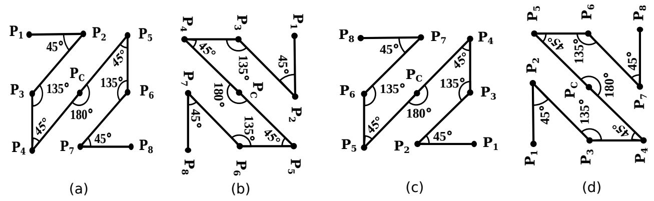

Fig. 1 *Variants of Local ZigZag structure (3 × 3), where (a)-(d) represent of four different (S^1, S^2, S^3, S^4) orders.*

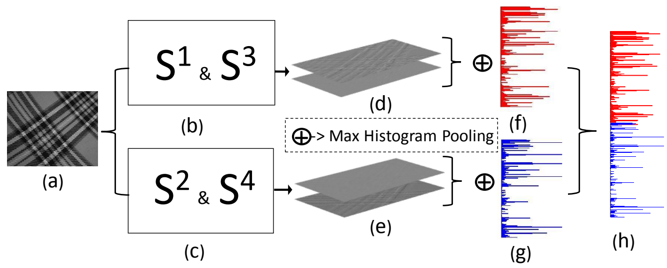

Fig. 2 *Proposed Framework: (a) Texture image (b)-(c) Different local ZigZag structures are stacked to form two complementary* LZP *groups i.e, (S^1, S^3) and (S^2, S^4) where S^1, S^2, S^3, and S^4 are shown in Fig. 1 (d)-(e) Two complementary* LZP *maps (f)-(g) The* LZMH *is build from* CLZP *(h) Represents final* LZMHPP *histogram.*

The LZP code generates 2^N distinct decimal values from the sign of the differences between center and its neighborhoods and generally described by the string of N-bit binary, where 2^{n-1} represents the weights of the n-th bit. Since, we calculate LZP by the thresholding within a 3×3 window , the LZP output ranges between 0 to 255. Due to the visibility of angular relationship from second to the seventh pixel of the ZigZag structure of Fig. 1, the two neighbors of each pixel extends the angle by either 45° or 180° or 135°. While in LBP, the angular relation between any two successive sample point with respect to its reference pixel is always to be a constant (45° for 8 sample points) and shows no visually perceptible angular relationship between two alternate pixel with respect to their intermediate. These fluctuations of LZP angles among the sample points allow to capture more frequent changes in local micro texture pattern. This micro structure forms four different orders of Local ZigZag kernel based on their starting point shown in Fig. 1(a)-(d), and which may make LZP as a better rotational invariant texture descriptor than traditional descriptors.

First, the four different LZP maps are extracted from the input texture image I (Fig. 2(a)) using the four different orders ZigZag structure so, called S^1, S^2, S^3 and S^4 (Fig. 2(b)-(c)). Then, the extracted LZP maps are used to formed two groups of complementary local ZigZag Pattern (CLZP) i.e. CLZP_1 and CLZP_2 by stacking the output of two complementary ZigZag structure together as shown in Fig. 2(d)-(e) and represented as,

$$\begin{aligned}\text{CLZP}_1 &= \{\text{LZP}(i,j)^{S1}, \text{LZP}(i,j)^{S3}\} \\ \text{CLZP}_2 &= \{\text{LZP}(i,j)^{S2}, \text{LZP}(i,j)^{S4}\}\end{aligned} \tag{3}$$

In this context, the complementary ZigZag structure indicates that it is the 90° rotated version of another one, which allows to encode the complementary rotational invariant texture information. After that, the local gray scale distribution of LZPs for each pixel (x, y) of the texture image I with dimension $M_x \times M_y$ is computed. It is indexed by building a discrete distribution of 2^N bins of Lzp codes, more formally the histogram is defined by,

$$\text{LZPH}(\lambda) = f(\text{CLZP}, S^{(g)}) \;\; S(g) = \begin{cases} (2g-1, 2g+1), & g == 1 \\ (g, 2g), & g == 2 \end{cases}$$

$$f(\text{LZP}(i,j), t) = \sum_{i=2}^{M_x-1} \sum_{j=2}^{M_y-1} \phi(\text{LZP}(i,j), t, \Omega), \; \phi(p,q) = \begin{cases} 1, & p == q \\ 0, & else \end{cases} \tag{4}$$

where $\Omega \in [0, \mathcal{L}]$, and $\mathcal{L}$ represents the maximum value of CLZP pattern.

The histogram codes of LZMH_1 and LZMH_2 for each reference bins are calculated using a max histogram pooling ($\oplus$) operation between two

This letter is published by IET, Electronics Letters. The copyright has been transferred to IET. The final letter DOI: https://doi.org/10.1049/el.2018.7631

Table 1: Texture classification performance on different texture descriptors

Dataset	Classifier	Accuracy		Feature Length
		KTH-TIPS	CUReT	
LBP	SVM	97.75	98.10	256
LTP		97.87	98.02	512
LDP		97.92	98.08	1024
LTrP		97.90	98.00	767
LDZP		98.00	98.17	354
LZMHPP		98.75	98.38	512
LZMHPP	MLP	98.30	98.20	512

Table 2: SVM accuracy of LZMHPP with different ZigZag combinations.

Descriptors	KTH-TIPS	CUReT
LZMHPP$^{(S1,S2)(S3,S4)}$	97.50 ± 0.824	97.41 ± 0.612
LZMHPP$^{(S1,S4)(S2,S3)}$	97.30 ± 0.560	97.50 ± 0.470
LZMHPP$^{(S1,S3)(S2,S4)}$	98.75 ± 0.621	98.38 ± 0.524

LZPH and represented as (see Fig. 2(f)-(g)),

$$\mathrm{LZMHP}_1 = \mathrm{LZPH}(i,j)^{S1} \oplus \mathrm{LZPH}(i,j)^{S3}$$
$$\mathrm{LZMHP}_2 = \mathrm{LZPH}(i,j)^{S2} \oplus \mathrm{LZPH}(i,j)^{S4} \quad (5)$$

The final LZMHPP descriptor as shown in Fig. 2(h) is formed by concatenating the two codes generated by LZMHP_1 and LZMHP_2 as,

$$\mathrm{LZMHPP} = \{\mathrm{LZMHP}_1, \mathrm{LZMHP}_2\} \quad (6)$$

The LZMHPP is efficient as it only doubles the dimension of basic LZP.

Experimental Results and Discussions: To inspect the texture classification performance of the proposed LZMHP descriptor, the experiments are carried out on two commonly used well-known publicly available texture databases, namely KTH-TIPS [8] and CUReT [9], respectively. The **CUReT database** [9] [10] is designed to contain large intra-class variation and is widely used to estimate the classification performance. It contains 61 texture classes with 92 images per class, which are cropped into 200×200 regions and converted to gray scale. The images are captured under different illumination and viewing directions with constant scale. The **KTH-TIPS** database [8] is extended by introducing new samples of 10 CUReT textures. It contains texture images with 3 different poses, 4 illuminations, and 9 different scales of size 200×200 and where each class contains 81 samples.

The performance of proposed LZMHPP descriptor is evaluated using the parametric multiclass support vector machine (LIBSVM [11]) classifier. The classification performance is computed as the average of K-fold ($K = 10$) cross validation. The estimated classification accuracy of proposed LZMHPP descriptor is compared with the several state-of-the-art texture descriptors, like, LBP [1], LTP [2], LDP [3], LTrP [4], and LDZP [7] in Table 1. The LZMHPP has better classification rates for both the texture databases compared to the existing texture descriptors. This is because LZMHPP encodes the local ZigZag relationship and extracts the encoded features using max pooling strategy. The discriminative ability is increased by encoding the local relation through multiple ZigZag fashion. The robustness of descriptor is increased by applying the max pooling operation. It is evident that the proposed LZMHPP descriptor outperforms other existing methods and achieves classification rates of 98.25% and 98.50% for KTH-TIPS and CUReT dataset, respectively. In addition to SVM classifier the performance of the proposed LZMHPP descriptor is also evaluated using multilayer perceptron (MLP) and achieves comparable classification rates 98.30% and 98.20% for KTH-TIPS and CUReT dataset, respectively. It shows that SVM outperforms MLP in the proposed framework. Table 1 also shows the length of the different feature vectors including proposed one. It can be observed that the length of the proposed LZMHPP descriptor is lower or comparable to LTP, LDP, and LTrP descriptors.

In the proposed framework (shown in Fig. 2), the complementary LZP maps are stacked together. We performed an experiment to find out the best suitable pairs together to maximize the utilization of complementary information. There are three different ways, we can choose the ZigZag structure to form two independent groups taken from a total 4 available structures (i.e., S^1, S^2, S^3, and S^4) (Fig. 1). We have used following three possible

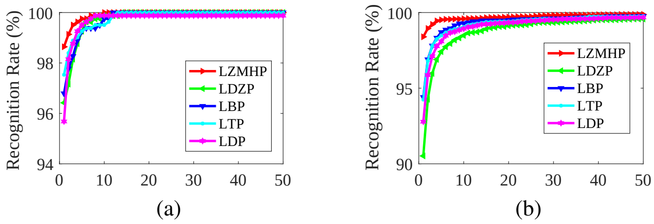

Fig. 3 *Cumulative Matching Characteristic (CMC) Curve for classification performance of descriptors on (a) KTH-TIPS and (b) CUReT databases.*

combinations, LZMHPP$^{(S1,S2)(S3,S4)}$, LZMHPP$^{(S1,S4)(S2,S3)}$, and LZMHPP$^{(S1,S3)(S2,S4)}$. The classification performance for 3 different LZMHPP descriptors are shown in Table 2. It can be observed from this result that among three possible combinations, the LZMHPP$^{(S1,S3)(S2,S4)}$ descriptor outperforms others. It shows that the groups $(S1, S3)$ and $(S2, S4)$ have the highest complementary information in its LZP maps. The performance of the proposed LZMHPP descriptor is also compared with existing descriptors using Cumulative Matching Characteristic (CMC) curve shown in Fig. 3. The CMC curves also show that the proposed LZMHPP descriptor outperforms the state-of-the-art descriptors.

Conclusion: This letter proposed a new texture descriptor named as Local ZigZag Max Histograms of Pooling Pattern (LZMHPP) by exploiting the max pooling over the histograms of features of several local ZigZag relationships. The utilization of multiple ZigZag features boosts the discriminative power and max pooling of histograms increases the robustness of the proposed descriptor against geometrical changes. The proposed descriptor is tested over two benchmark texture datasets and the results are compared with state-of-the-art descriptors. Very promising performance is observed for proposed LZMHPP descriptor.

Swalpa Kumar Roy, and Bidyut B. Chaudhuri (*Computer Vision & Pattern Recognition Unit, Indian Statistical Institute, Kolkata - 700108, India*) E-mail: swalpa@cse.jgec.ac.in; bbc@isical.ac.in

Shiv Ram Dubey (*Computer Vision Group, Indian Institute of Information Technology, Sri City, Andhra Pradesh - 517646, India*) E-mail: srdubey@iiits.in

References

1 Ojala, T., Pietikainen, M., Maenpaa, T.: 'Multiresolution gray-scale and rotation invariant texture classification with local binary patterns', *IEEE TPAMI*, 2002, **24**, (7), pp. 971–987

2 Tan, X., Triggs, B.: 'Enhanced local texture feature sets for face recognition under difficult lighting conditions', *IEEE TIP*, 2010, **19**, (6), pp. 1635–1650

3 Zhang, B., Gao, Y., Zhao, S., Liu, J.: 'Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor', *IEEE TIP*, 2010, **19**, (2), pp. 533–544

4 Murala, S., Maheshwari, R., Balasubramanian, R.: 'Local tetra patterns: a new feature descriptor for content-based image retrieval', *IEEE TIP*, 2012, **21**, (5), pp. 2874–2886

5 Dubey, S.R., Singh, S.K., Singh, R.K.: 'Novel local bit-plane dissimilarity pattern for computed tomography image retrieval', *Electronics Letters*, 2016, **52**, (15), pp. 1290–1292

6 Roy, S.K., Chanda, B., Chaudhuri, B.B., Ghosh, D.K., Dubey, S.R. 'A complete dual-cross pattern for unconstrained texture classification'. In: 4th IAPR ACPR. (IEEE, 2017. pp. 741–746

7 Roy, S.K., Chanda, B., Chaudhuri, B.B., Banerjee, S., Ghosh, D.K., Dubey, S.R.: 'Local Directional ZigZag Pattern: A Rotation Invariant Descriptor for Texture Classification', *Pattern Recognition Letters*, 2018, **108**, pp. 23–30

8 Hayman, E., Caputo, B., Fritz, M., Eklundh, J.O. 'On the significance of real-world conditions for material classification'. In: ECCV. (Springer, 2004. pp. 253–266

9 Dana, K.J., Van.Ginneken, B., Nayar, S.K., Koenderink, J.J.: 'Reflectance and texture of real-world surfaces', *ACM TOG*, 1999, **18**, (1), pp. 1–34

10 Zhang, J., Marszałek, M., Lazebnik, S., Schmid, C.: 'Local features and kernels for classification of texture and object categories: A comprehensive study', *IJCV*, 2007, **73**, (2), pp. 213–238

11 Chang, C.C., Lin, C.J.: 'Libsvm: a library for support vector machines', *ACM transactions on intelligent systems and technology (TIST)*, 2011, **2**, (3), pp. 27