An Integrated Descriptor for Texture Classification

Vu-Lam Nguyen, Ngoc-Son Vu, Hai-Hong Phan, Philippe-Henri Gosselin

An Integrated Descriptor for Texture Classification

Vu-Lam Nguyen Ngoc-Son Vu Hai-Hong Phan Philippe-Henri Gosselin
ETIS - ENSEA/ Université de Cergy-Pontoise/ CNRS UMR 8051
95000-Cergy, France
Email: "lam.nguyen,son.vu,thi-hai-hong.phan,philippe-henri.gosselin"
{ @ensea.fr }

***Abstract*—Regarding texture features, Local-based methods such as Local Binary Pattern (LBP) and its variants are computationally efficient high-performing but sensitive to noise, and suffering global structure information loss. By contrast, filter-based counterparts, the Scattering Transform for instance, are tolerant to noise and translation but often lack of small local structure information. In this paper we propose an integration of those to take full advantages of both local and global features. In this way, LBP is used for extracting local features while the Scattering Transform feature plays the role of a global descriptor. In addition to the combination of these two state-of-the-art features, we further integrate a new preprocessing technique called biologically-inspired filtering (BF) as well as an efficient PCA classifier. Intensive experiments conducted on many texture benchmarks such as CUReT, UIUC, KTH-TIPS2b, and OUTEX show that our combined method not only outweighs each one which stands alone but also competes with state-of-the-art on the experimented datasets.**

***Index Terms* - Image texture, Image classification, Image texture analysis.**

1. Introduction

In the recent decades, computer vision researchers have been witnessing a lot of researches drawn on texture because it is the fundamental appearance element of materials or objects. Main problems consist of segmentation, shape from texture, synthesis, and classification. Also, texture is a crucial clue for distinguishing things in the real world. For example, animals are able to be discriminated depending upon the texture of their skin, the category of soil, sand, rock and so forth can be differentiated base on the texture of their surfaces. Texture classification is a research field that categorizes image data into more readily interpretable information, which is able to be used in a wide range of applications such as industrial inspection, image retrieval [1], medical imaging, remote sensing [2], object recognition, and facial recognition [3], [4], [5]. For this reason, it is a very inspiring subject.

Texture images can be represented for classification by using two main methods. Large scale filter-based approaches through the statistical distributions of their responses, take early works in [6], [7], [8], [9], [10] and more recent ones in [11], [12], [13], [14] for examples. On the other hand, texture classification relies on features extracted from small scale neighborhoods using the pixel intensities, some early research of this direction includes works in [15], [16] while their more sophisticated up-to-date variants can be seen in [17], [18], [19], [20]. If the former method attracts a more global structure information of images, the latter demonstrates that a good discrimination is able to be achieved through exploiting the distributions of pixel neighborhoods.

Among the local neighborhood-based techniques, LBP [15] has drawn considerable attention since its proposal. The LBP family has already been used in many other applications such as image retrieval, and facial image analysis [5] because it is not only simple to implement, real time running but is also a highly distinguishable descriptor. However, the conventional version has some limitations, such as small spatial support region, loss of global textural information, also sensitivity to noise. Many LBP variants were proposed to overcome those, Completed LBP (CLBP) [17] is one of them. These variants still struggle to get high performance on image datasets with variant in scale, translation, and deformation. On the other hand, the Scattering Transform introduced in Scattering NetWork (ScatNet) by Mallat et al. [12], which applies Wavelet Transform in a deep convolution network, copes well with those characteristics of data. However, these Scattering Transform features do not capture well the small local structure information.

In this paper, we propose to use the strength of both LBP and Scattering transform in a "Hybrid" descriptor for texture classification, CLBP [17] is used instead of the original version because of its higher performance. A preprocessing algorithm called BF as well as an efficient PCA classifier are also used. The rest of paper is organized as follows. After discussing about related work in Section 2, the proposed approach is introduced in Section 3. Section 4 presents experimental results, and conclusions are drawn in Section 5.

2. Related Work

In this section, LBP, CLBP, Scattering Transform, BF, and PCA classifier are reviewed.

2.1. Brief Overview of the LBP and CLBP

The LBP method, proposed by Ojala et al. [15], encodes the pixel-wise information in images. LBP encoding is:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1)$$

where g_c represents the gray value of the center pixel whereas g_p $(p = 0, ..., P-1)$ denotes the gray value of the neighbor pixel on a circle of radius R, and P is the total number of the neighbors. A given texture image is then represented by a histogram of LBP codes. Ojala et al. also introduced a rotation invariant complement called uniform patterns LBP^{riu2} which have less than two "one-to-zero or vice versa" transitions.

Guo et al. [17] recently proposed CLBP descriptors, by which image local differences are decomposed into two complementary components, the signs (s_p) and the magnitudes (m_p): $s_p = s(g_p - g_c)$, $m_p = |g_p - g_c|$ where g_p, g_c and $s(x)$ are defined as in (1). Two operators called CLBP-Sign ($CLBP\_S$) and CLBP-Magnitude ($CLBP\_M$), respectively, are proposed to encode them, where the $CLBP\_S$ is equivalent to the conventional LBP, and the $CLBP\_M$ measures the local variance of magnitude. The $CLBP\_M$ is defined as follows:

$$CLBP\_M_{P,R} = \sum_{p=1}^{P-1} t(m_p, c)2^p, \; t(x, c) = \begin{cases} 1, & x \geq c \\ 0, & x < c \end{cases}$$

where threshold c is the mean value of m_p of the whole image. CLBP-Center ($CLBP\_C$) operator extracts the local central information as $CLBP\_C_{P,R} = t(g_c, c_I)$ where threshold c_I is set as the average gray level of the whole image. Over all descriptor is obtained by combining the three operators $CLBP\_S$, $CLBP\_M$ and $CLBP\_C$.

2.2. Review of Scattering Transform

The Scattering transform [11] is calculated by using Scattering Network (ScatNet). ScatNet is actually a deep convolution network, in which the Wavelet Transformation followed by modulus non-linearities operators are consecutively computed. As illustrated in the second component of (Figure 1), S_0x and U_1x are generated by the first wavelet modulus operator W_1 based on input image x.

$W_1(x) = (S_0(x), U_1(x))$,
$S_0x(u) = x \star \phi_j(u) = \sum_v x(v)\phi_j(u - v)$,
$U_1(x) = x \star \psi_{j,\theta}(x)$,
$\phi_j(u) = 2^{-2j}\phi(2^{-j}u)$ is a Gaussian low pass filter. This leads to the averaged image S_0x being almost invariant to rotations and translations up to 2^j pixels. While it loses the high frequencies of x, these will be gotten back by the convolution with high pass wavelet filters. Then the wavelet ψ is rotated by θ angles and dilated by 2^j in order to achieve rotation covariant coefficients.

$S(x)$ is called the scattering coefficient of the network, $U(x)$ is the wavelet modulus coefficient, and $\star$ is a convolution operator. Finally, the scattering features vector of an image is obtained by concatenating all scattering coefficients of the network, $S(x) = (S_0(x), S_1(x), S_2(x))$.

2.3. Biologically Inspired Filtering

Biologically Inspired Filtering (BF) [21] imitates the human retina mechanism to extract more detail information of a given image when being used as a preprocessing step. It enhances performance of different features in terms of discriminative power for texture classification, including CLBP.

In general, BF consists of two steps, as follows (Figure 3):

Step 1: given an image I_{in}, it is first filtered by a band-pass Difference of Gaussians (DoG). $I_{bf} = DoG \star I_{in}$, where $\star$ is the convolution operator, and

$$DoG = \frac{1}{2\pi\sigma_1^2} e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \frac{1}{2\pi\sigma_2^2} e^{-\frac{x^2+y^2}{2\sigma_2^2}}.$$

σ_1, σ_2 are the standard deviations of the low pass filters.

Step 2: the filter responses are then decomposed into two "maps" corresponding to the image details alongside two sides of the image edge:

$$I_{bf}^{+} = \begin{cases} I_{bf}(p) & if \quad I_{bf}(p) \geq \epsilon \\ 0 & otherwise \end{cases}$$
$$I_{bf}^{-} = \begin{cases} |I_{bf}(p)| & if \quad I_{bf}(p) \leq -\epsilon \\ 0 & otherwise \end{cases}.$$

The term bf refers to "Biologically-inspired Filtering", p is the considered pixel, ϵ is a threshold. In features extraction step, instead of using the input image I_{in}, features are first extracted from two images, I_{bf}^{-} and I_{bf}^{+}, and are then combined together.

2.4. PCA Classifier

A generative classifier called Principal Component Analysis (PCA) [13] was proved to have decent performance for ScatNet in case of small training dataset. PCA Classifier is described as following.

Given a test image X, $\widetilde{S}X$ denotes the scattering transfrom of X and its dilated version $\mathfrak{D}_j X$.

$$\widetilde{S}X = \left(\sum_{0 \leq j < H} 1 \right)^{-1} \sum_{0 \leq j < H} \widetilde{S}\mathfrak{D}_j X. \quad (2)$$

The representation of $\widetilde{S}X$ used at test time is therefore a Scattering Transform.

Let $P_{U_c}\widetilde{S}X$ denotes the orthogonal projection of $\widetilde{S}X$ in the scattering space U_c of a given class c. The principal components space U_c is approximately computed from the singular value decomposition (SVD) of the matrix of centered training sample $\widetilde{S}\mathfrak{D}_j X_{c,i} - \mu_c$ with all possible samples i dilated by 2^j for a given class c. The PCA classification computes the class $\hat{c}(X)$ base on the minimum

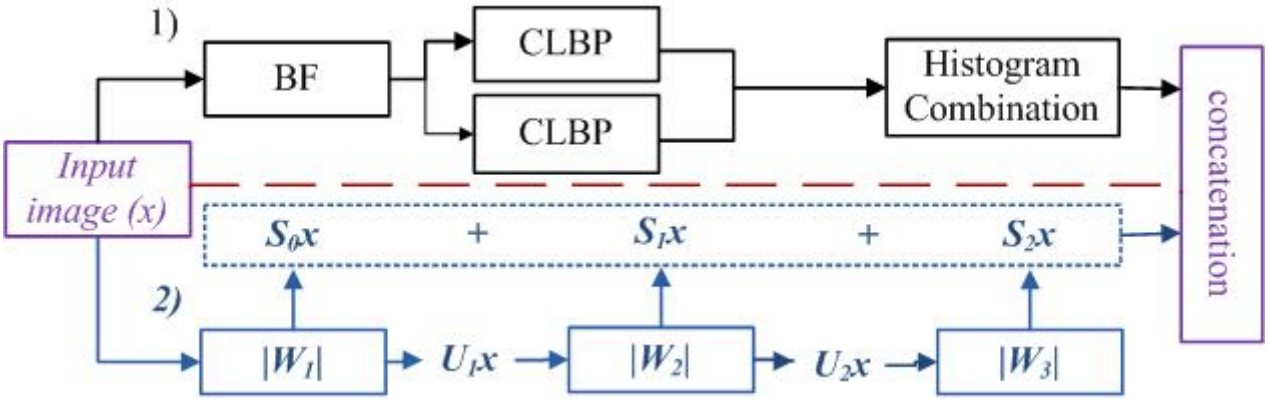

Figure 1. Integrated Descriptor consists of two components separated by the dashed line:
1) BF takes one image as input, and generates two outputs I_{bf}^{+} and I_{bf}^{-} which are further used as CLBP's inputs, namely BF+CLBP component.
2) ScatNet Component: scattering representation is computed by a cascade of wavelet-modulus operators $|W_m|$. Every $|W_m|$ has one input and two outputs which are the scattering coefficients $S_m x$ and the next layer wavelet modulus coefficients $U_{m+1}x$. The latter is used for further transformation. Final features are the concatenation of those generated by the components.

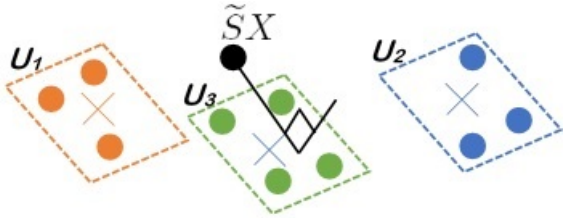

Figure 2. PCA-classifier classifies a test image X based on the minimum distance from Scattering Transform $\widetilde{S}X$ to subspace $\mu_c + U_c$.

distance $\left\|(Id - P_{V_c})(\widetilde{S}X - \mu_c)\right\|$ from $\widetilde{S}X$ to the space $\mu_c + U_c$, (Figure 2)

$$\hat{c}(X) = arg\min_{c} \left\|(Id - P_{U_c})(\widetilde{S}X - \mu_c)\right\|^2 \qquad (3)$$

where μ_c is the average of Scattering Transform $\widetilde{S}X_{c,i}$ for all training samples $X_{c,i}$ of class c, and

$$\mu_c = \left(\sum_{0\le j<H} 1\right)^{-1} \sum_{0\le j<H} \widetilde{S}\mathfrak{D}_j X_{c,i} \qquad (4)$$

The value H quantifies the ranges of scale invariance, e.g for a training set with dilated versions of each $X_{c,i}$ by different scaling factors 2^j for $0 \le j < H$. H is the range of scale invariance, limited by the size of images. Typically, $H = 2$ and sample j at half integer which leads to 4 scaling factors $\{1, \sqrt{2}, 2, 2\sqrt{2}\}$, and the dilated samples $\mathfrak{D}_{c,i}(u) = X_{c,i}(2^j u)$.

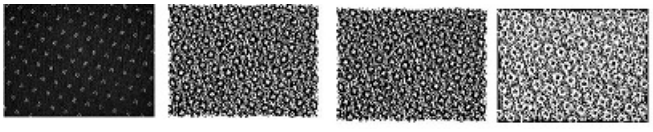

Figure 3. BF processing chain: I_{in}, I_{bf} , I_{bf}^{+} , I_{bf}^{-} (left to right).

3. Proposed Method

LBP proposed by Ojala et al. [15] and its variants such as CLBP [17] can capture well the small local structure information of the image data. On the other hand, the Scattering Transform method [11], [12], [13] extracts a wider range of signals for its features. According to our observation of those, we intuitively think that if there is a descriptor which employs both small local and global structure information it would make texture classification have promising results.

It can be vividly seen that conventional LBP features are vulnerable to noise because the most important part of this method is based on the center pixel threshold. A measure to this is that we make use of our previous works, BF preprocessing technique [21], before extracting LBP features. Scattering Transform method also has its weak point, a lack of small local structure information due to its feature dimension reduction strategy, downsampling. This can be compensated for using LBP features. We do not utilize the BF preprocessing technique for ScatNet because this network contains band pass filters whose functionality are similar to those of BF. The overall idea is illustrated in Figure 1.

For a given input image x, we first apply BF+CLBP and the Scattering Transform on the image simultaneously, then aggregate their outputs (features) to form final features. These features capture well both local and global structure information of texture images. Specifically, BF splits the input image x into I_{bf}^{+} and I_{bf}^{-} (Figure 3), BF+CLBP representation of image x is the combination of CLBP features of I_{bf}^{+} and I_{bf}^{-}, namely BF+CLBP features. Similarly, the ScatNet features are gained by an average value of the concatenation of scattering coefficients (S_0x, S_1x, S_2x). Finally, an aggregation of those forms our integrated features. We also discover that the PCA classifier, which is used for ScatNet in [11], [12], [13], has a higher performance than nearest neighbor classifier for our descriptor in term of classification accuracy.

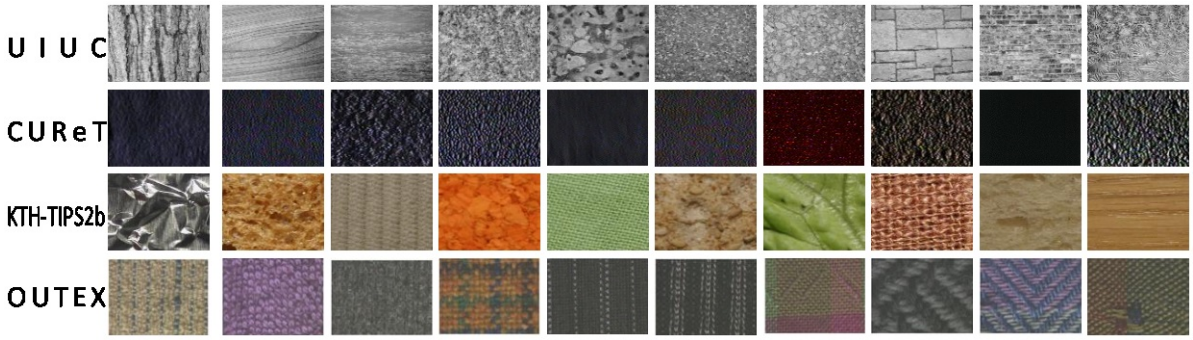

Figure 4. Images from 4 different datasets: UIUC,CUReT, KTH-TIPS2b, and OUTEX

It should be noticed that we focus mainly upon building a descriptor based on the Scattering Transform rather than techniques of Scatnet [12] such as multi-scale average, and multi-scale training to augment classification results.

TABLE 1. INTEGRATED DESCRIPTOR WITH BF+CLBP+SCATNET SINGLE SCALE: CLASSIFICATION ACCURACY(%) AND COMPARING WITH THOSE OF WELL-KNOWN METHODS ON UIUC DATASET

Method	Accuracy(%)
BF+CLBP_S/M	96.48
BF+CLBP_S/M/C	96.28
ScatNet	95.20
BF+CLBP_S/M+ScatNet	**98.63**
BF+CLBP_S/M/C+ScatNet	**98.69**
Zhang et al. [24]	98.70
scLBP [18]	98.45
SRP [19]	98.40
VZ-joint [22]	97.83
Hayman et al. [23]	92.00

TABLE 2. INTEGRATED DESCRIPTOR WITH BF+CLBP+SCATNET: CLASSIFICATION ACCURACY(%) ON CURET DATASET

Method	Single-scale(%)	Multi-scale(%)
BF+CLBP_S/M	95.66	98.60
BF+CLBP_S/M/C	95.84	98.56
ScatNet	99.10	99.10
BF+CLBP_S/M+ScatNet	99.28	99.53
BF+CLBP_S/M/C+ScatNet	99.29	99.51

TABLE 3. COMPARING PROPOSAL CLASSIFICATION ACCURACY(%) WITH THOSE OF WELL-KNOWN METHODS ON CURET

Method	Accuracy(%)
AnID (ours)	**99.53**
scLBP [18]	99.29
BRINT [27]	99.27
Broadhurst [26]	99.22
SRP [19]	99.05
MRELBP [20]	99.02
Hayman et al. [23]	98.46
VZ-MR8 [25]	97.79
VZ-joint [22]	97.66
Zhang et al. [24]	95.30

TABLE 4. INTEGRATED DESCRIPTOR WITH BF+CLBP+SCATNET MULTI-SCALE: PROPOSAL CLASSIFICATION ACCURACY(%) ON KTHTIPS2B DATASET AND COMPARING WITH THOSE OF WELL-KNOWN METHODS

Method	Accuracy(%)
BF+CLBP_S/M	67.68
BF+CLBP_S/M/C	68.31
ScatNet	60.40
BF+CLBP_S/M+ScatNet	**78.09**
BF+CLBP_S/M/C+ScatNet	**77.71**
MRELBP [20]	77.91
ELBP [28]	64.84
LTP [3]	62.12
PRICoLBP [29]	61.17
LBP [15]	60.35
VZ-Path [22]	60.07
VZ-MR8 [25]	55.7

4. Experimental Validation

4.1. Experimental Settings

We analyze the effectiveness of our method by doing experiments on four popular texture databases and follow their testing protocols. In order to have reliable results, we run the PCA classifier [13] 100 times on each dataset by randomly choosing the training and testing sets, with an exception for outex and KTH-TIPS-2b which have a predefined sets of those. Then the average results of all splits are reported. In addition to the BF preprocessing technique [21], PCA classifier [13], and Scaterring Network (ScatNet) [11], [31] which plays Scattering Transform role are used in our experiment.

Parameters for the BF preprocessing technique are: $\sigma_1 = 1.25, \sigma_2 = 6, \epsilon = 0.15$ as recommended in [21].

Regarding CLBP [17], the parameters chosen as follows: the number of neighbors (P=16), and Radius (r=1,2,3) or (r=3,5,7) for three different scales. Arguments of ScatNet [11], [31] for the Scattering Transform are selected such that number of scale is J=4 for datasets with resolution of images at 200×200 or below, J=5 otherwise, Orientations of filter bank (L=8), number of ScatNet Orders (M=3). Sifre et al. showed in [12] that the maximum number of layers of the network is 3, if

TABLE 5. INTEGRATED DESCRIPTOR WITH BF+CLBP+SCATNET: CLASSIFICATION ACCURACY(%) ON OUTEX DATASET

Methods	OUTEX TC10		OUTEX TC12_00		OUTEX TC12_01	
	Single-scale	Multi-scale	Single-scale	Multi-scale	Single-scale	Multi-scale
BF+CLBP_S/M	99.40	99.60	95.05	97.29	96.34	97.99
BF+CLBP_S/M/C	99.43	99.50	96.88	97.13	97.38	98.40
ScatNet	98.39	98.39	96.23	96.23	98.38	98.38
BF+CLBP_S/M+ScatNet	99.77	99.82	97.91	98.22	98.40	99.40
BF+CLBP_S/M/C+ScatNet	99.66	99.87	98.15	98.43	98.77	99.63

TABLE 6. COMPARING PROPOSAL CLASSIFICATION ACCURACY(%) WITH THOSE OF WELL-KNOWN METHODS ON OUTEX DATASET

Method	TC10	TC12	
		t184	horizon
MRELBP [20]	99.87	99.49	99.75
AnID(ours)	**99.87**	**98.43**	**99.63**
LTP [3]	98.54	92.59	89.17
LBP [15]	97.70	87.30	86.40
PRICoLBP [29]	94.48	92.57	92.50
VZ-MR8 [25]	93.59	92.55	92.82
NRLBP [30]	93.44	86.13	87.38
VZ-Path [22]	92.00	91.41	92.06

this number exceeds 3 then the energy will decay, and so no more useful signal for discrimination.

Experiments were conducted on following datasets, samples of those represented in Figure 4.

UIUC [32] has 25 classes of texture, each class contains 40 different images which has resolution of 640×480 include changes of viewing angle, scale, and illumination conditions. The mean classification accuracy, 100 random splits between training and testing with a half of samples per class chosen for training, is reported.

CUReT [33] database contains 61 texture classes, 205 images per class, acquired at different viewpoints, illumination, and orientations. There are 118 images shot from a viewing angle of less than 60 degrees. We choose a subset 92 images from 118 each class for our experiments, in total $61 \times 92 = 5612$ images are selected. Large images are cropped to (200×200) across all texture classes. All cropped regions are converted to grey scale. For the experiments on CUReT database, we follow the common training and testing scheme used in [18], [20], [21], a half of class samples chosen for training while the remaining for testing. Splits are implemented 100 times independently, the average accuracy over 100 randomly partitions are reported.

The material databases **KTHTIPS2b** [34], with 3 viewing angles, 4 illuminants, and 9 different scales, produce 432 images per class, with the image size of 200×200 and 11 classes in total. Regarding the KTH-TIPS2b databases, we follow the testing and training protocols applied in [20], [35]. Only unseen data is used for testing, with three out of four samples used for training and the remaining for testing.

OUTEX [36] database contains textural images which are captured from a wide variety of real material surfaces. We consider the two commonly used test suites, Outex_TC_00010 (TC10) and Outex_TC_00012 (TC12), containing 24 classes with up to 200 texture images per class. This database is built by taking images under three different illuminations (”horizon”, ”inca”, and ”t184”) with resolution of images at 128×128. Standard protocols with predefined training and testing sets are exploited for this dataset.

4.2. Classification Results

Intensive experiments were conducted on four texture datasets, the results are compared with well-known and state-of-the-art of those, we chose the highest results reported by relevant articles for the comparison. Our method reaches state-of-the-art on UIUC [32], CUReT [33], KTHTIPS2-b [34], and are comparable on three testing suites of Outex [36]. Details are shown in Table 1,3,4,6. The bold lines are results of our proposal while others are taken from works in [18], [20]. It is worth noticing that our novel descriptor is consistently better than the ones it inherits from on the experimented datasets while we do not use the multi-scale training technique of ScatNet [12].

Details about classified accuracy on experimented databases as:

UIUC: We get around 2% classification enhancement on this dataset when using our single scale integrated descriptor (from 96.48% to 98.69%) comparing to the its original version. Results from Table 1 shows that our accurate rate on this dataset is just slightly lower than those of works in [24] while beating all others. Data used for comparing in talbe 1 is taken from works in [18].

CUReT: The correct classified rate of our method reaches State-of-the-art on this dataset, the proportion are above 99% which gains a competitive advantage over the results that have been reported in research by Ryu et al. [18].

KTHTIPS2b: As can be seen from Table 4, our method outperforms BF+CLBP and Scattering Transform descriptor up to 8%. Also, our classification result on this database is the best among those of recent state-of-the-art reported in [20].

OUTEX: In case of Outex database, experiment conducted on 2 test suites, Outex_TC10 and Outex_TC12. For Outex_TC10, the accuracy is at 99.87% (exactly the same result of recent state-of-the-art reported in [20]). Results on Outex_TC12 ”horizon” and ”t184” are also very good, they are at 98.43% and 99.63% respectively. Ours outperforms all competing results except MRELBP [20]. Details can be seen in (Table 6).

5. Conclusion

In this paper, we have proposed a three-in-one handcrafted descriptor, namely an integrated descriptor. It takes full advantages of the BF preprocessing technique (our previous works [21]), the local LBP features, and the global ones extracted from ScatNet. It is proved by experiments that this novel descriptor enhances distinctiveness of texture while preserving the robustness to variations in illumination, rotation, and noise. Overall, ScatNet and LBP are not concurrent, but complementary while the preprocessing technique makes the descriptor more robust. Future study can be drawn on the same domain with scale variation tolerance by using multi-scale training technique.

References

[1] C. Schmid, "Constructing models for content-based image retrieval," in *2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001)*, 2001.

[2] A. H. Schistad and A. K. Jain, "Texture analysis in the presence of speckle noise," in *Geoscience and Remote Sensing Symposium, 1992. IGARSS'92. International*, vol. 2. IEEE, 1992, pp. 884–886.

[3] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *Analysis and Modeling of Faces and Gestures*. Springer, 2007, pp. 168–182.

[4] N. Vu, H. M. Dee, and A. Caplier, "Face recognition using the POEM descriptor," *Pattern Recognition*, vol. 45, no. 7, pp. 2478–2488, 2012.

[5] N.-S. Vu, "Exploring patterns of gradient orientations and magnitudes for face recognition," *IEEE Trans. Information Forensics and Security*, vol. 8, no. 2, pp. 295–304, 2013.

[6] O. G. Cula and K. J. Dana, "Compact representation of bidirectional texture functions," in *2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001)*, 2001.

[7] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 1, pp. 55–73, 1990.

[8] A. Laine and J. Fan, "Texture classification by wavelet packet signatures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1186–1191, 1993.

[9] D. Charalampidis and T. Kasparis, "Wavelet-based rotational invariant roughness features for texture classification and segmentation," *IEEE Trans. Image Processing*, vol. 11, no. 8, pp. 825–837, 2002.

[10] M. Varma and A. Zisserman, "Classifying images of materials: Achieving viewpoint and illumination independence," in *European Conference on Computer Vision*. Springer-Verlag, 2002, pp. 255–271.

[11] S. Mallat, "Group Invariant Scattering," *Communications on Pure and Applied Mathematics*, vol. 65, no. 10, pp. 1331–1398, 2012.

[12] L. Sifre and S. Mallat, "Rotation, scaling and deformation invariant scattering for texture discrimination," in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 1233–1240.

[13] J. Bruna and S. Mallat, "Invariant scattering convolution networks," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 8, pp. 1872–1886, 2013.

[14] L. Sifre and S. Mallat, "Rigid-motion scattering for texture classification," *CoRR*, vol. abs/1403.1687, 2014.

[15] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2002.

[16] M. Varma and A. Zisserman, "Texture classification: Are filter banks necessary?" in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2003.

[17] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Trans. Image Processing*, vol. 19, no. 6, pp. 1657–1663, Jun. 2010.

[18] J. Ryu, S. Hong, and H. S. Yang, "Sorted consecutive local binary pattern for texture classification," *IEEE Transactions on Image Processing*, vol. 24, no. 7, pp. 2254–2265, 2015.

[19] L. Liu, P. W. Fieguth, D. Hu, Y. Wei, and G. Kuang, "Fusing Sorted Random Projections for Robust Texture and Material Classification," vol. 25, no. 3, 2015.

[20] L. Liu, P. W. Fieguth, M. Pietikäinen, and S. Lao, "Median robust extended local binary pattern for texture classification," in *2015 IEEE, ICIP 2015*, 2015, pp. 2319–2323.

[21] N. Vu, T. P. Nguyen, and C. Garcia, "Improving texture categorization with biologically-inspired filtering," *Image Vision Comput.*, vol. 32, no. 6-7, pp. 424–436, 2014.

[22] M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 2032–2047, 2009.

[23] E. Hayman, B. Caputo, M. Fritz, and J. Eklundh, "On the significance of real-world conditions for material classification," in *ECCV*, 2004.

[24] Y. Xu, X. Yang, H. Ling, and H. Ji, "A new texture descriptor using multifractal analysis in multi-orientation wavelet pyramid," in *CVPR*, 2010, pp. 161–168.

[25] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images," *Int. Journal of Computer Vision*, vol. 62, no. 1-2, pp. 61–81, 2005.

[26] R. E. Broadhurst, "Statistical estimation of histogram variation for texture classification," in *Proc. Intl. Workshop on texture analysis and synthesis*, 2005, pp. 25–30.

[27] L. Liu, B. Yang, P. Fieguth, Z. Yang, and Y. Wei, "Brint: A binary rotation invariant and noise tolerant texture descriptor," in *International Conference on Image Processing*, Melbourne, 2013.

[28] L. Liu, L. Zhao, Y. Long, G. Kuang, and P. Fieguth, "Extended local binary patterns for texture classification," *Image Vision Comput.*, vol. 30, no. 2, pp. 86–99, Feb. 2012.

[29] X. Qi, R. Xiao, C.-G. Li, Y. Qiao, J. Guo, and X. Tang, "Pairwise rotation invariant co-occurrence local binary pattern," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, no. 11, pp. 2199–2213, 2014.

[30] J. Ren, X. Jiang, and J. Yuan, "Noise-resistant local binary pattern with an embedded error-correction mechanism," *IEEE Trans. Image Processing*, vol. 22, no. 10, pp. 4049–4060, 2013.

[31] J. Bruna and S. Mallat, "Invariant scattering convolution networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1872–1886, 2013.

[32] S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine regions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1265–1278, 2005.

[33] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real-world surfaces," *ACM Trans. Graph.*, vol. 18, no. 1, pp. 1–34, 1999.

[34] B. Caputo, E. Hayman, M. Fritz, and J.-O. Eklundh, "Classifying materials in the real world," *Image Vision Comput.*, vol. 28, no. 1, pp. 150–163, 2010.

[35] B. Caputo, E. Hayman, and P. Mallikarjuna, "Class-specific material categorisation," in *10th IEEE International Conference on Computer Vision (ICCV 2005)*, 2005.

[36] T. Ojala, T. Menp, M. Pietikinen, J. Viertola, J. Kyllnen, and S. Huovinen, "Outex - new framework for empirical evaluation of texture analysis algorithms," in *16th Int. Conf. on Pattern Recognition*, 2002, pp. 701–706.