

Statistical binary patterns for rotational invariant texture classification

Thanh Phuong Nguyen, Ngoc-Son Vu, Antoine Manzanera

▶ **To cite this version:**

Thanh Phuong Nguyen, Ngoc-Son Vu, Antoine Manzanera. Statistical binary patterns for rotational invariant texture classification. Neurocomputing, 2016, 10.1016/j.neucom.2015.09.029 . hal-01245103

HAL Id: hal-01245103
https://hal.science/hal-01245103

Submitted on 16 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Statistical binary patterns for rotational invariant texture classification

Thanh Phuong Nguyen[a,b,d,*], Ngoc-Son Vu[c], Antoine Manzanera[d]

[a] *Aix Marseille Université, CNRS, ENSAM, LSIS, UMR 7296, 13397 Marseille, France.*
[b] *Université de Toulon, CNRS, LSIS, UMR 7296, 83957 La Garde, France*
[c] *ETIS, ENSEA, Université de Cergy-Pontoise, 6 Avenue du Ponceau, 95014 Cergy-Pontoise CEDEX, France*
[d] *U2IS-Robotics & Vision, ENSTA-Paristech, Université de Paris-Saclay, 828 Bd des Maréchaux, 91762 Palaiseau CEDEX, France*

Abstract

A new texture representation framework called statistical binary patterns (SBP) is presented. It consists in applying rotation invariant local binary pattern operators (LBP^{riu2}) to a series of moment images, defined by local statistics uniformly computed using a given spatial support. It can be seen as a generalisation of the commonly used complementation approach (CLBP), since it extends the local description not only to local contrast information, but to higher order local variations. In short, SBPs aim at expanding LBP self-similarity operator from the local gray level to the regional distribution level. Thanks to a richer local description, the SBPs have better discrimination power than other LBP variants. Furthermore, thanks to the regularisation effect of the statistical moments, the SBP descriptors show better noise robustness than classical CLBPs. The interest of the approach is validated through a large experimental study performed on five

*Corresponding author
Email address: thanh-phuong.nguyen@univ-tln.fr (Thanh Phuong Nguyen)

texture databases: KTH-TIPS, KTH-TIPS 2b, CUReT, UIUC and DTD. The results show that, for the four first datasets, the SBPs are comparable or outperform the recent state-of-the-art methods, even using small support for the LBP operator, and using limited size spatial support for the computation of the local statistics.
Keywords: texture classification, local binary pattern, statistical moments

1. Introduction

Texture analysis is an active research topic in computer vision and image processing. It has been widely used in many applications including medical image analysis, remote sensing, object recognition, document classification, content-based image retrieval and many more. As one of the major problems in texture analysis, texture classification has received considerable attention over the two last decades, and many novel methods have been proposed. [1, 2, 3, 4, 5].

The texture classification problem is typically divided into the two subproblems of representation and classification. It is generally agreed that texture features play a very important role. If inadequate feature descriptors are used, even the best classifier will fail to achieve good results. Despite decades of research efforts on texture description, it remains a challenging problem, because of many variations that can affect texture patterns, like viewpoint changes, illumination variation, rotation, noise, and so on. Thus, a good texture descriptor should not only capture highly discriminative information but also be robust to extrinsic changes.

Most earlier texture representations focused on filter banks and the sta-

tistical distributions of their responses. Among the popular descriptors in these approaches are Gabor filters [6], MR8 [7], Leung and Malik's filters [8], steerable filters [9] or wavelets [10]. More recently, Ojala et al. with the Local Binary Patterns (LBP) [1] and Varma and Zisserman with the VZ-Joint classifier [11] have shown that local intensities or differences in a small patch can produce better performance than filter banks with large spatial support. Due to their great computational efficiency and good texture characterization, the LBPs are often preferred to other frameworks. They have been applied in many applications of computer vision and a large number of LBP variants [12, 13, 14, 15, 16, 17, 18] have been introduced.

Although significant progress has been made, most LBP variants still have prominent drawbacks, like noise sensitivity [12], and a limitation to the contrast information. In this paper, we propose a generic statistical approach, which considerably enhances both the discriminative power of LBPs and their robustness to small variations, notably image noise.

The proposed algorithm encodes the local structure of the gray level distribution, summarised by moments of order up to 4. We first consider a series of local statistical moments. The self-similarity LBP operators are applied on those images. We further encode a binary image, obtained by thresholding the moment images with the corresponding global averages. Integrating these complementary ingredients, our algorithm encodes rich descriptive information while providing robustness to environmental changes. Experimental results on five large and representative texture databases show that our approach favourably competes with state-of-the-art descriptors.

Our method is naturally related to filter bank approaches [6, 7, 8, 9],

and can also be seen as a pre-processing technique, since the moment images are the output of local filters. By the way, moment images have been partially used before, in [19, 20, 21, 1] where mean or variance filters are used for noise robustness or for complementary information purposes. In this paper, we show that the series of moment images calculted from a pre-defined structuring element at different orders (up to 4^{th} order) bring more textural information since the regional gray level distribution is better represented using different statistical moments.

The rest of this paper is organised as follows. Section 2 discusses the related literature on LBP-based texture descriptors, and Section 3 details the proposed approach. Section 4 develops the implementations being evaluated and their parameter setting. Experimental results are presented in Section 5 and conclusions are finally drawn in Section 6.

2. Related work

Basic LBPs are "micro" features, capturing the relationships between pixels in small-scale neighbourhoods, but they have several limitations, such as small spatial support region, loss of local textural information, rotation and noise sensitivities. To overcome these drawbacks, a lot of efforts have been made. To recover from the loss of information, local image contrast was introduced by Ojala et al. [1] as a complementary measure, and better performance has been reported. In a "completed" LBP model, Guo et al. [13] included both the magnitudes of local differences and the pixel intensity itself, and improved again the performance. In terms of locality, Liao et al. [12] proposed to extract global features from the Gabor filter responses as

a complementary descriptor. Liu et al. [17] have recently generalized LBPs with two complementary types of features extracted from local patches, based on pixel intensities and differences, reporting impressive texture classification rates. In [16], an "LBP like" feature, the Local Binary Count (LBC), is proposed, by encoding only the number of neighbours whose value is higher than the centre pixel. Heikkila et al. [22] exploited circular symmetric LBP (CS-LBP) for local interest region description. Zhao et al. [23] presented the LBP histogram Fourier (LBP-HF) features while Maani [24] defined local frequency components from the 1D Fourier transform of the neighbouring function. In [25], Chen et al. proposed the WLD (Weber Local Descriptor), a bio-inspired extension of the LBP based on Weber's law. For preprocessing step, Gabor filters [26] are widely used for capturing more global information. Different neighbourhoods have also been used: elliptical [27], three-patch or four-patch neighbourhood [28] to exploit anisotropic information. Logically, multi-scale or multi-structure approaches [29, 21] were considered to gather information at larger scales. Zhao et al. [30] combined LBP with covariance matrix to improve the performance.

2.1. LBP

The LBP descriptor, first proposed by Ojala et al. [1], encodes the spatial relations in images. Let f be a discrete image, modelled as a mapping from $\mathbb{Z}^2$ to $\mathbb{R}$. The original LBP encoding of f is defined as the following mapping from $\mathbb{Z}^2$ to $\{0,1\}^P$:

$$\mathrm{LBP}_{P,R}(f)(\mathbf{z}) = \left(s(f(\mathbf{y}_p) - f(\mathbf{z}))\right)_{0\leq p<P}, \tag{1}$$

$$\text{with } s(x) = \begin{cases} 1, x \geq 0 \\ 0, \text{otherwise.} \end{cases}$$

Here $\mathbf{y}_p$ $(0 \leq p < P)$ are the P neighbours of pixel $\mathbf{z}$, whose values are evenly measured (or interpolated) on the circle of radius R centred on $\mathbf{z}$.

The uniformity measure of an LBP is defined as follows:

$$U(\text{LBP}_{P,R}) = \sum_{p=1}^{P} |\text{LBP}_{P,R}^{p} - \text{LBP}_{P,R}^{p-1}|, \tag{2}$$

where $\text{LBP}_{P,R}^{p}$ is the p-th bit of $\text{LBP}_{P,R}$, and $\text{LBP}_{P,R}^{P} = \text{LBP}_{P,R}^{0}$. An LBP is called uniform if $U(\text{LBP}_{P,R}) \leq 2$. Ojala et al. observed that, on natural texture images, most patterns are uniform. Finally the rotation invariant uniform LBP is defined as follows:

$$\text{LBP}_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} \text{LBP}_{P,R}^{p}, \text{if } U(\text{LBP}_{P,R}) \leq 2 \\ P + 1, \text{otherwise.} \end{cases} \tag{3}$$

LBP^{riu2} proved [1] a very efficient local texture descriptor and then has been intensively used in texture classification. Uniform patterns are considered as more reliable and more statistically significant. Furthermore, ignoring non-uniform patterns considerably reduces the length of the descriptor, with only $P + 2$ distinct $LBP_{P,R}^{riu2}$ compared to 2^P distinct $LBP_{P,R}$.

2.2. Complemented LBP

Guo et al. [13] presented a state-of-the-art variant by considering the local differences as two complementary components, signs: $s_p = s(f(\mathbf{y}_p) - f(\mathbf{z}))$ and magnitudes: $m_p = |f(\mathbf{y}_p) - f(\mathbf{z})|$. They proposed two operators, called

CLBP-Sign ($CLBP_S$) and CLBP-Magnitude ($CLBP_M$) to code these two components. The first operator is identical to the LBP. The second one which measures the local variance of magnitude is defined as follows:

$$\text{CLBP_}M_{P,R}(f)(\mathbf{z}) = (s(m_p - \tilde{m}))_{0 \leq p < P}\,, \tag{4}$$

where $\tilde{m}$ is the mean value of m_p for the whole image. In addition, Guo et al. observed that the local value itself carries important information. Therefore, they defined the operator CLBP-Center (CLBP_C) as follows:

$$\text{CLBP_}C(f)(\mathbf{z}) = s(f(\mathbf{z}) - \tilde{f}), \tag{5}$$

where $\tilde{f}$ is set as the mean gray level of the whole image. Because these operators are complementary, their combination leads to a significant improvement in texture classification, then this variant is also considered as a reference LBP method.

2.3. Noise tolerant LBP variants

To address the problem of noise sensitivity, different methods have been proposed, that can be generally categorised into three groups.

The most popular approach uses different symbols to code small differences. Tan and Triggs [31] proposed local ternary patterns (LTP), using ternary $\{-1, 0, 1\}$ instead of binary encoding, where 0 is used when the pixel difference is less than a given threshold. Several variants have been presented: Nanni et al. [32] proposed quinary encoding, while Akhloufi and Bendada [33] determined the threshold adaptively; Liao et al. [34] used pixel ratio instead of differences in order to deal with complex gray scale intensity changes.

The second approach uses pre-processing filters in encoding step to reduce the influence of noise. Instead of using central pixel as threshold, several authors used the median [35] or the mean value of neighbouring pixels [19, 20, 21]. Ojala et al. [1] also used variance as complementary information to make LBP more robust to noise. Recently, Liu et al. [36] presented a mechanism of "Average Before Quantization" in encoding step. Those methods make the texture representations more robust against noise, while allowing them to capture more spatial information.

The last one is based on an error-correction mechanism or soft assignment to recover uncertain bit in LBPs. Ren et al. [37] presented a mechanism to recover the corrupted patterns by encoding small pixel difference as an uncertain bit that is determined later based on other bits. Iakovidis et al. [38] incorporated fuzzy logic in LBP encoding to cope with uncertainty in noisy ultrasound images.

3. Statistical Binary Pattern

The Statistical Binary Pattern (SBP) representation aims at enhancing the expressiveness and discrimination power of LBPs for texture modelling and recognition, while reducing their sensitivity to unsignificant variations (e.g. noise). The principle consists in applying rotation invariant uniform LBP to a set of images corresponding to local statistical moments associated to a spatial support. The resulting code forms the Statistical Binary Patterns (SBP). Then a texture is represented by joint or marginal distributions of SBPs. The classification can then be performed using nearest neighbour criterion on classical histogram metrics like χ^2. We now detail those different

steps.

3.1. Moment images

A real valued 2d discrete image f is modelled as a mapping from $\mathbb{Z}^2$ to $\mathbb{R}$. The spatial support used to calculate the local statistics is modelled as $B \subset \mathbb{Z}^2$, such that $O \in B$, where O is the origin of $\mathbb{Z}^2$.

The r-order moment image associated to f and B is also a mapping from $\mathbb{Z}^2$ to $\mathbb{R}$, defined as:

$$m^r_{(f,B)}(\mathbf{z}) = \frac{1}{|B|} \sum_{\mathbf{b} \in B} \left(f(\mathbf{z} + \mathbf{b})\right)^r \tag{6}$$

where $\mathbf{z}$ is a pixel from $\mathbb{Z}^2$, and $|B|$ is the cardinality of the structuring element B. Accordingly, the r-order centred moment image ($r > 1$) is defined as:

$$\mu^r_{(f,B)}(\mathbf{z}) = \frac{1}{|B|} \sum_{\mathbf{b} \in B} \left(f(\mathbf{z} + \mathbf{b}) - m^1_{(f,B)}(\mathbf{z})\right)^r \tag{7}$$

where $m^1_{(f,B)}(\mathbf{z})$ is the average value (1-order moment) calculated around $\mathbf{z}$. Finally the r-order normalised centred moment image ($r > 2$) is defined as:

$$\beta^r_{(f,B)}(\mathbf{z}) = \frac{1}{|B|} \sum_{\mathbf{b} \in B} \left(\frac{f(\mathbf{z} + \mathbf{b}) - m^1_{(f,B)}(\mathbf{z})}{\sqrt{\mu^2_{(f,B)}(\mathbf{z})}} \right)^r \tag{8}$$

where $\mu^2_{(f,B)}(\mathbf{z})$ is the variance (2-order centred moment) calculated around $\mathbf{z}$.

3.2. Statistical Binary Patterns

Let R and P denote respectively the radius of the neighbourhood circle and the number of values sampled on the circle. For each moment image M, one statistical binary pattern is formed as follows:

- one $(P+2)$-valued pattern corresponding to the rotation invariant uniform LBP coding of M:

$$\mathrm{SBP}_{P,R}(M)(\mathbf{z}) = \mathrm{LBP}^{riu2}_{P,R}(M)(\mathbf{z}) \tag{9}$$

- one binary value corresponding to the comparison of the centre value with the mean value of M:

$$\mathrm{SBP}_{C}(M)(\mathbf{z}) = s(M(\mathbf{z}) - \tilde{M}) \tag{10}$$

Where s denotes the pre-defined sign function, and $\tilde{M}$ the mean value of the moment M on the whole image. $\mathrm{SBP}_{P,R}(M)$ then represents the structure of moment M with respect to a local reference (the centre pixel), and $\mathrm{SBP}_{C}(M)$ complements the information with the relative value of the centre pixel with respect to a global reference ($\tilde{M}$). As a result of this first step, a $2(P+2)$-valued scalar descriptor is then computed for every pixel of each moment image.

3.3. Texture Descriptors

Let $\{M_i\}_{1\leq i\leq n_M}$ be the set of n_M computed moment images. $\mathrm{SBP}^{\{M_i\}}$ is defined as a vector valued image, with n_M components such that for every $\mathbf{z} \in \mathbb{Z}^2$, and for every i, $\mathrm{SBP}^{\{M_i\}}(\mathbf{z})_i$ is a value between 0 and $2(P+2)$.

If the image f contains a texture, the descriptor associated to f is made by the histogram of values of $\mathrm{SBP}^{\{M_i\}}$. We consider two kinds of histograms. The first one is the joint histogram H defined as follows:

$$H : [\![0\,;2(P+2)[\![^{n_M} \to \mathbb{N}$$

$$H(\mathrm{v}) = |\{\mathbf{z}; \mathrm{SBP}^{\{M_i\}}(\mathbf{z}) = \mathrm{v}\}| \tag{11}$$

Depending on the size of the texture images, the joint distribution may become too sparse when the dimension (i.e. the number of moments) increase. Then we also consider the marginal histograms $\{h_i\}_{i \leq n_M}$ defined as:

$$h_i : [\![0\,; 2(P+2)[\![\to \mathbb{N}$$

$$h_i(n) = |\{\mathbf{z}; \mathrm{SBP}^{\{M_i\}}(\mathbf{z})_i = n\}| \tag{12}$$

The texture descriptor can then be formed by a joint histogram H or by the concatenation of n_M marginal histograms $\{h_i\}$. The length of the descriptor vector is $[2(P+2)]^{n_M}$ in the first case, and $2n_M(P+2)$ in the second case.

3.4. Texture Classification

Two texture images being characterised by their respective histogram (descriptor) F and G, their texture dissimilarity metrics is calculated using the classical χ^2 distance between distributions:

$$\chi^2(F, G) = \sum_{i=1}^{d} \frac{(F_i - G_i)^2}{F_i + G_i} \tag{13}$$

where d is the number of bins (dimension of the descriptors). Our classification is then based on a nearest neighbour criterion. Every texture class with label λ is characterised by a prototype descriptor K_λ, with $\lambda \in \Lambda$. For an unknown texture image f, its descriptor D_f is calculated, and the texture label is attributed as follows:

$$l(f) = \arg\min_{\lambda \in \Lambda} \chi^2(D_f, K_\lambda) \tag{14}$$

4. Implementations

In this section, we detail the instances of the SBP framework that will be evaluated in this paper, as well as their parameter settings.

4.1. Mean and Variance

Our first experiments focused on $\mathrm{SBP}_2(\mathrm{P,R}) = \mathrm{SBP}_{P,R}^{m_1\mu_2}$, i.e., the SBPs obtained with the mean m_1 and the variance μ_2 [1]. Using two orders of moments, the size of the joint histogram in texture descriptor remains reasonable. Figure 1 illustrates the calculation of the texture descriptor using m_1 and μ_2 images. The spatial relationships between local structures on each image are characterised using $\mathrm{LBP}_{P,R}^{riu2}$. In addition, two binary images are computed by thresholding the moment images with respect to their average values. For each moment, the local pattern then have $2(P+2)$ distinct values. Finally, the joint histogram of the two local descriptors is used as the texture feature and is denoted $\mathrm{SBP}^{m_1\mu_2}$. Therefore, the feature vector length is $4(P+2)^2$.

On the histogram depicted in Figure 1, the peak lines correspond to non-uniform bins, which mainly appear due to quantisation noise on constant valued area, and then are fully representative of the texture. Aside from the peaks, the histogram shows a relatively smooth structure, thanks to the LBP^{riu2} coding which possesses, unlike traditional LBP coding, a natural metrics. The complementarity of the two moment components is visible on the histograms. Furthermore, the global shape of the histogram is representative of the texture class, as can be seen on Figure 2, where Figures 2(b)

[1] It was also presented in the preliminary paper [39].

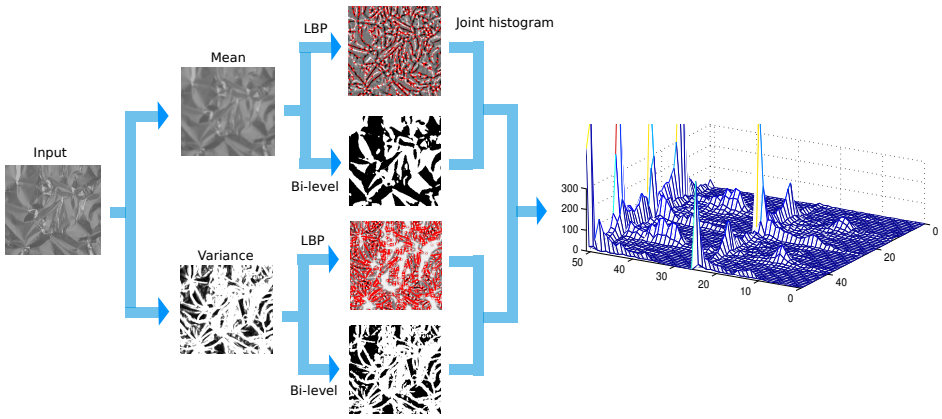


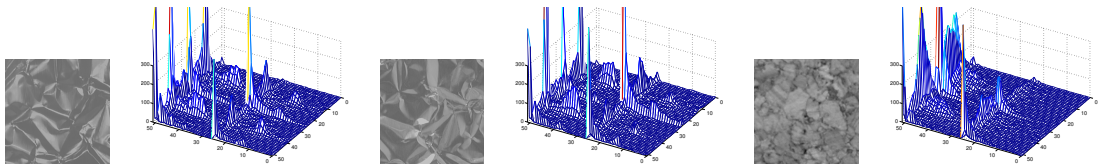
Figure 1: Texture representation using SBP_2 =$\text{SBP}^{m_1\mu_2}$. The structuring element $B = \{(1,5),(2,8)\}$ (see Section 4.4) is used for computing the moment images, and $\text{LBP}^{riu2}_{24,3}$ is used for LBP encoding (red pixels correspond to non-uniform patterns).

and (d) show SBP_2 descriptors for two images from the same texture class, whereas Figure 2(f) shows the SBP_2 descriptor for another texture. Figures 3(a)-(d) explicits the interpretation of the SBP_2 descriptor, according to the meaning of the 4 regions within the 2d histogram:

- First quadrant (Figure 3(a)): Homogeneous dark areas
- Second quadrant (Figure 3(b)): Dark areas near contours
- Third quadrant (Figure 3(c)): Homogeneous light areas
- Fourth quadrant (Figure 3(d)): Light areas near contours

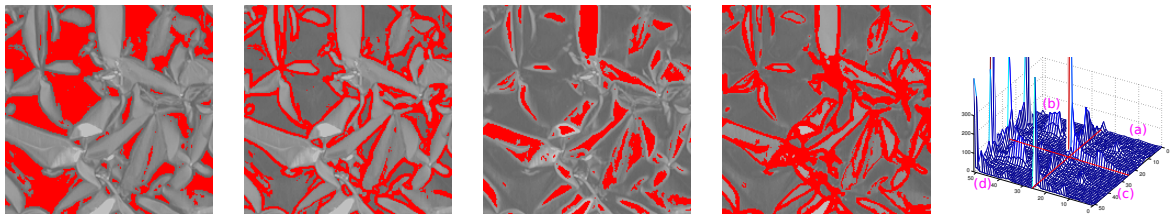
Several remarks can be made on the properties of $\text{SBP}^{m_1\mu_2}$ and its link to existing work.

- *Robustness to noise:* m_1 and μ_2 act like a pre-processing step which reduces small local variations and then enhances the significance of the



(a) (b) (c) (d) (e) (f)

Figure 2: (a)-(b), (c)-(d) and (e)-(f): Texture images and their corresponding SBP_2 descriptors.



(a) (b) (c) (d) (e)

Figure 3: Interpretation of the SBP_2 descriptor (e) and its 4 regions (a-b) with, on the left, the corresponding back-projected pixels.

binary pattern with respect to the raw images.

- *Rotation invariance:* isotropic structuring elements which discards all orientation information should be used in order to keep the rotation invariance property of the local descriptor.

- *Information richness:* moment images already convey information about local structures, and thus, applying LBPs at a larger-scale on those moment images allows our algorithm to capture more global information. Also, the two orders of moments provide complementary information of the spatial structure.

There are also links between $\text{SBP}^{m_1\mu_2}$ and the CLBP descriptors of Guo et al. [13] (see Sec. 2). First, the binary images used in SBP correspond to CLBP_C operator. Second, the respective roles of m_1 and μ_2 are somewhat similar to the CLBP_S and CLBP_M operators. However in CLBP and its variants, CLBC [16] and CRLBP [20], the magnitude component (CLBP_M) is more a contrast information complementing CLBP_S, whereas in our algorithm, SBP^{μ_2} represents the local structure of a contrast map, and can be considered independently on SBP^{m_1}.

4.2. Higher order moments

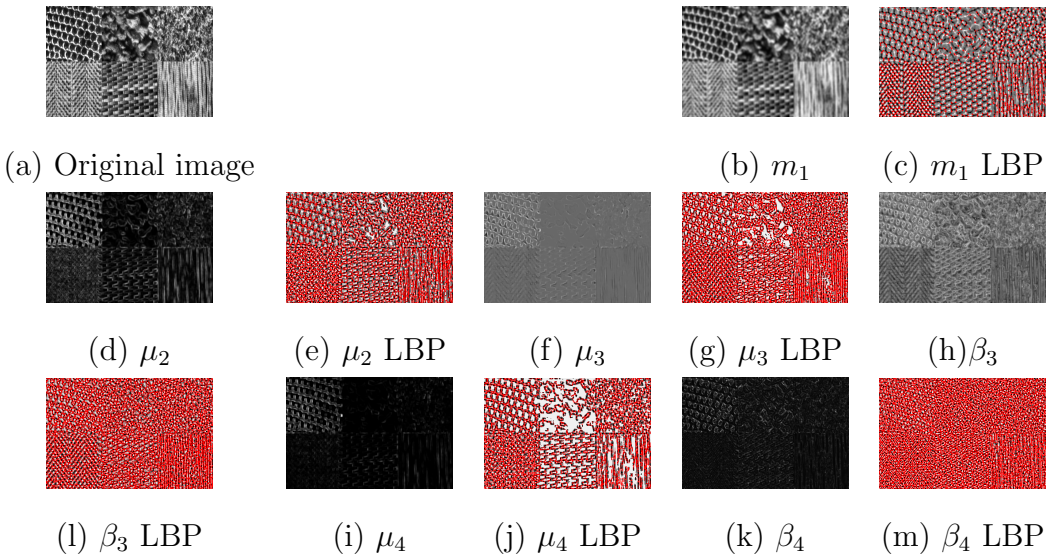
We have also evaluated the SBP model on higher order moments. Indeed, the objective of the SBP framework is to extend the LBP texture descriptors from the local level, represented by pixel $\mathbf{z}$, to the regional distribution level of $\mathbf{z} + B$, by approximating the distribution to a set of statistical moments. However the mean and variance only describe faithfully the distribution in

particular cases like Gaussian distributions, which is obviously insufficient in many locations of natural texture images. Higher order moments may capture relevant information to complete the description. For example, skewness (β_3) is a measure of (dis)symmetry of the distribution while kurtosis (β_4) describes the height and sharpness of its peak.

Figure 4 shows different moment images for different orders, and their LBP encoding ($(P, R) = (24, 3)$) where red pixels correspond to non-uniform patterns. The different patterns appearing in the moment images and their LBP illustrate the complementarity of the different orders of moments. At the same time, the increasing number of red pixels in the LBP^{riu2} images as the order grows, is an illustration of the instability of the higher order moments: because they amplify the small variations, more unstable local variations appear on the higher order moment images. The consequence is a decrease in the number of significant LBP values in the images. Thus, we limited the evaluation to the 4^{th} order moment.

Regarding the size of the texture descriptor, it will naturally grow as the number of moments increases. When using joint histograms, the descriptor size is $(2(P+2))^n$, where P is the number of neighbours used in LBP, and n is the number of moment images. When using marginal histograms, the size is only $2n(P+2)$, but at the price of a significant loss of information. We then propose a trade-off based on concatenation of joint histograms corresponding to pairs of moment images.

Formally, we can recursively define the higher order SBP hybrid texture descriptor as follows. Let M_1 and M_2 be moments or combinations of moments by their joint or concatenated histogram. We shall denote $SBP^{M_1 M_2}$



(a) Original image (b) m_1 (c) m_1 LBP

(d) μ_2 (e) μ_2 LBP (f) μ_3 (g) μ_3 LBP (h)β_3

(l) β_3 LBP (i) μ_4 (j) μ_4 LBP (k) β_4 (m) β_4 LBP

Figure 4: Several moment images and their LBP encodings at different orders of a textured image. $\text{LBP}^{riu2}_{24,3}$ is used and non-uniform patterns are highlighted by red pixels in images of LBP encoding.

(resp. $SBP^{M_1_M_2}$) the texture descriptor made by the joint (resp. concatenated) histograms constructed from SBP^{M_1} and SBP^{M_2}.

In our experiments for higher order moments, we have only considered pairs of moments for joint histograms. Specifically, for the third order moments, we have evaluated the two following descriptors (their size is $3(2(P+2))^2$ in both cases):

- *Non normalised third order*
  $\text{SBP}'_3(P,R) = \text{SBP}^{m_1\mu_2_m_1\mu_3_\mu_2\mu_3}_{P,R}$
- *Normalised third order*
  $\text{SBP}_3(P,R) = \text{SBP}^{m_1\mu_2_m_1\beta_3_\mu_2\beta_3}_{P,R}$

Accordingly, for the fourth order, we have evaluated the two following

descriptors (their size is $6(2(P+2))^2$ in both cases):

- *Non normalised fourth order*
  $\mathrm{SBP}_4^{'}(P,R) = \mathrm{SBP}_{P,R}^{m_1\mu_2_m_1\mu_3_\mu_2\mu_3_m_1\mu_4_\mu_2\mu_4_\mu_3\mu_4}$

- *Normalised fourth order*
  $\mathrm{SBP}_4(P,R) = \mathrm{SBP}_{P,R}^{m_1\mu_2_m_1\beta_3_\mu_2\beta_3_m_1\beta_4_\mu_2\beta_4_\beta_3\beta_4}$

4.3. Multiscale complementary SBP descriptor

Our SBP descriptors can be easily extended by considering other aspects of LBP encoding. We address hereafter two popular extensions proposed by other LBP-based variants for improvement of SBP descriptors.

- Multiscale approach [29]
- Complementary information [13]

For multiscale approach, several supporting neighbourhoods $\{(P_1,R_1),(P_2,R_2),\ldots,(P_k,R_k)\}$ are considered instead of one single scale (P,R). The multiscale SBP descriptors (M_SBP) are obtained by concatenating SBPs calculated from each single support.

- $\mathrm{M_SBP}_2 = \mathrm{SBP}_2(P_1,R_1)_\ldots_\mathrm{SBP}_2(P_k,R_k)$
- $\mathrm{M_SBP}_3 = \mathrm{SBP}_3(P_1,R_1)_\ldots_\mathrm{SBP}_3(P_k,R_k)$
- $\mathrm{M_SBP}_4 = \mathrm{SBP}_4(P_1,R_1)_\ldots_\mathrm{SBP}_4(P_k,R_k)$

Concerning the second extension, complementary magnitude information is exploited by CLBP_M [13]. It is well-known that this kind of complementary information are very rich and it is the principal factor for improving the

performance in most recent LBP-based variants such as CLBP [13], CLBC [16], CRLBP [20], BRINT [36], disCLBP [40], SCLBP [41]. In this paper, the complementary descriptors are constructed in the same way as our SBP descriptors using CLBP_M operator.

- $\mathrm{SBP}_2^{\mathrm{M}} = (\mathrm{CLBP_M})^{m_1 m_2}$
- $\mathrm{SBP}_3^{\mathrm{M}} = (\mathrm{CLBP_M})^{m_1\mu_2_m_1\beta_3_\mu_2\beta_3}$
- $\mathrm{SBP}_4^{\mathrm{M}} = (\mathrm{CLBP_M})^{m_1\mu_2_m_1\beta_3_\mu_2\beta_3_m_1\beta_4_\mu_2\beta_4_\beta_3\beta_4}$

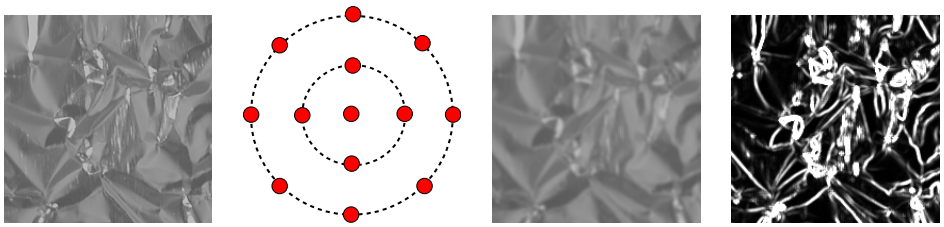
They are then concatenated with SBP descriptors to obtain complementary descriptors (C_SBP).

- $\mathrm{C_SBP}_2 = \mathrm{SBP}_2_\mathrm{SBP}_2^{\mathrm{M}}$
- $\mathrm{C_SBP}_3 = \mathrm{SBP}_3_\mathrm{SBP}_3^{\mathrm{M}}$
- $\mathrm{C_SBP}_4 = \mathrm{SBP}_4_\mathrm{SBP}_4^{\mathrm{M}}$

Finally these two last improvements (M_SBP and C_SBP) can be combined by concatenating the two descriptors to obtain the multiscale complementary descriptor (MC_SBP). In Section 5, MC_SBP is evaluated using five different scales $(P, R) \in \{(8,1),(8,2),(8,3),(8,4),(8,5)\}$.

4.4. Parameter settings

We now describe the different parameters that can be adjusted in the framework, and the different settings chosen for the evaluation. Apart from the choice of moments and their combinations, as presented above, two parameters need to be set in the calculation of the SBP:



Input B Mean (m_1) Variance (μ_2)

Figure 5: Computation of moment images using structuring element $B = \{(1,4),(2,8)\}$.

- the spatial support B for calculating the local moments, also referred to as structuring element.
- the spatial support $\{P, R\}$ for calculating the LBP.

Although those two parameters are relatively independent, it can be said that B has to be sufficiently large to be statistically relevant, and that its size should be smaller or equal to the typical period of the texture. Regarding $\{P, R\}$, it is supposed to be very local, to represent micro-structures of the (moment) images.

As mentioned earlier, for rotation invariance purposes, we shall use isotropic structuring elements. To be compliant with the LBP representation, we have chosen to define the structuring elements as unions of discrete circles: $B = \{\{P_i, R_i\}\}_{i \in I}$, such that $(P_i)_i$ (resp. $(R_i)_i$) is an increasing series of neighbour numbers (resp. radii). As an example, Figure 5 shows the moment images using a structuring element $B = \{(1,4),(2,8)\}$.

Specifically, we have evaluated the different following structuring elements: $\{(1,6)\}$, $\{(1,8)\}$, $\{(1,4),(2,8)\}$, $\{(1,5),(2,6)\}$, $\{(1,5),(2,8)\}$, $\{(1,5),(2,10)\}$, $\{(1,6),(2,10)\}$ and $\{(1,6),(2,12)\}$. In the next section, we only show experiments using structuring element $\{(1,5),(2,8)\}$ due to its good results on the

different datasets.

Regarding $\{P, R\}$, the spatial support of the LBP, we have chosen, for comparison purposes with the previous methods, to evaluate the three usual settings used in the literature: $\{8, 1\}$, $\{16, 2\}$, and $\{24, 3\}$ (except for CUReT dataset, where $\{8, 1\}$, $\{16, 3\}$, and $\{24, 5\}$ are considered).

5. Experiments

We present hereafter a comparative evaluation of the proposed descriptors [2]. The following descriptors will be addressed: $\text{SBP}_2 = \text{SBP}_{P,R}^{m_1\mu_2}$, $\text{SBP}_3^{'} = \text{SBP}_{P,R}^{m_1\mu_2_m_1\mu_3_\mu_2\mu_3}$, $\text{SBP}_3 = \text{SBP}_{P,R}^{m_1\mu_2_m_1\beta_3_\mu_2\beta_3}$, $\text{SBP}_4^{'} = \text{SBP}_{P,R}^{m_1\mu_2_m_1\mu_3_\mu_2\mu_3_m_1\mu_4_\mu_2\mu_4_\mu_3\mu_4}$, $\text{SBP}_4 = \text{SBP}_{P,R}^{m_1\mu_2_m_1\beta_3_\mu_2\beta_3_m_1\beta_4_\mu_2\beta_4_\beta_3\beta_4}$. In order to equitably evaluate the proposed descriptors, they are compared, together with other state-of-the-art descriptors, using the same classification criterion, i.e. nearest neighbour with χ^2 distance, which is the most commonly used method in the related literature. However, some of those methods may have better results using other classifiers such as SVMs.

5.1. Databases and Experimental Protocols

The effectiveness of the proposed method is assessed by a series of experiments on five large and representative databases: CUReT [42], UIUC [2], KTH-TIPS [43], KTH-TIPS2b [44] and DTD [45].

The CUReT (Columbia-Utrecht Reflection and Texture) database contains 61 texture classes, each having 205 images acquired at different view-

[2]Our code is publicly accessible via this address: `https://github.com/nguyenthanhphuong/SBP`

points and illumination orientations. There are 118 images shot from a viewing angle smaller than 60°. Following [11, 13], we selected 92 images out of 118, from which a sufficiently large region could be cropped (200×200) for all the texture classes. All the cropped regions were converted to grey scale (see examples in Figure 6).

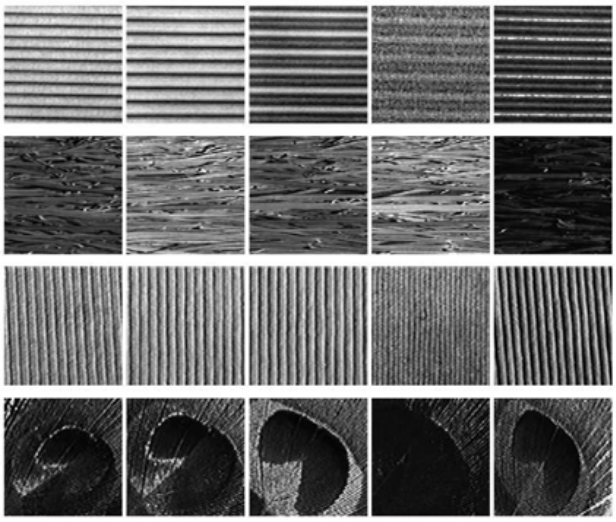


Figure 6: Examples of texture images from the CUReT database.

The UIUC texture database includes 25 classes with 40 images in each class. The resolution of each image is 640×480. The database contains materials imaged under significant viewpoint variations (see Figure 7).

The KTH-TIPS dataset [43] contains 10 materials. For each material, images were captured at nine scales spanning two octaves, viewed under three different illumination directions and three different poses, thus giving a total of 9 views per scale, and then 81 images per material. The result on KTH-TIPS database is reported as the mean classification rate over 100 random split into training and testing data, where in each material, 40 images

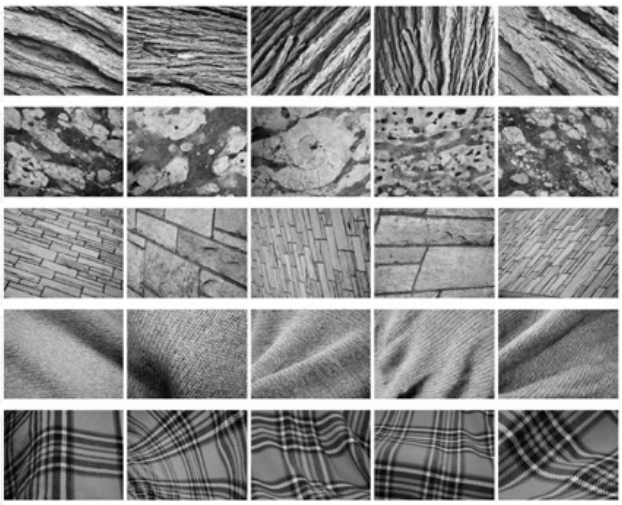


Figure 7: Examples of texture images on the UIUC database.

are used for training and the rest is used for testing.

The KTH-TIPS2b database contains images of 11 materials. Each material contains 4 physical samples taken at 9 different scales, 3 viewing angles and 4 illuminations, producing 432 images per class (see [44] for more details). Figure 8 shows examples of the 11 materials. All the images were cropped to 200 $\times$ 200 pixels and converted to grey scale. This database is considered more challenging than the previous version KTH-TIPS. In addition, it is more complete than KTH-TIPS2a where several samples have only 72 images.

The DTD database [45], consisting of 5640 images that were collected using Google and Flickr by entering perceptual attributes as search queries, contains 47 classes inspired from human perception, not material. Image sizes range between 300x300 and 640x640. Each class (category) has 120 images containing at least 90 % of the surface representing the category attribute.

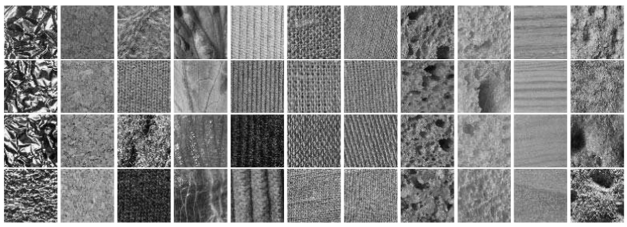


Figure 8: Examples of texture images on the KTHTIP2b database.

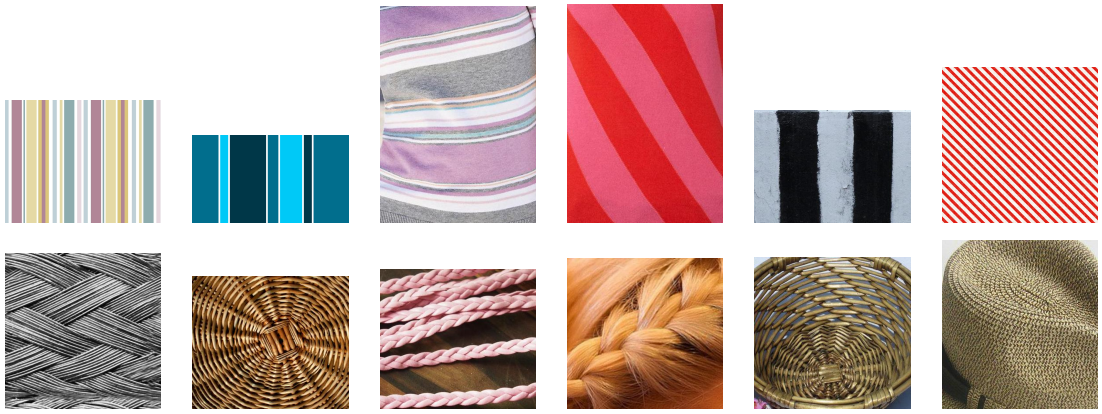


Figure 9: Examples of texture images on the DTD database. Each row contains images of a same category.

Figure 9 shows some texture images of this database.

5.2. Computational cost

We consider in this section the computational cost of the proposed descriptors, in terms of time and memory consumption, with respect to other LBP-based operators. Experiments on Outex [46] TC10 test suite containing 4 320 images of 128 $\times$ 128 pixels were performed on a machine with 2.0 GHz CPU, 4 Go RAM and Linux 3.2.0-23 kernel. Table 1 presents the computation time (in seconds) of different descriptors in three configurations of (P, R). We consider the structuring element $\{(1, 5), (2, 8)\}$ and report the

Table 1: Computational costs of different descriptors (FET: Feature Extraction Time, MT: Matching Time).

Descriptor	(P,R)=(8,1)			(P,R)=(16,2)			(P,R)=(24,3)		
	FET	Feature size	MT	FET	Feature size	MT	FET	Feature size	MT
$LBP^{riu2}_{P,R}$	52.71	10	1.18	98.43	18	2.13	145.07	26	3.40
CLBP_S/M/$C^{riu2}_{P,R}$	64.26	200	16.98	98.70	648	66.98	186.32	1152	456.98
$SBP^{m_1\mu_2}_{P,R}$ (SBP_2)	158.45	400	29.65	258.96	1296	136.34	511.78	2304	1044.74

total time for classifying the 3 840 test images against the 480 reference images. It can be seen from Table 1 that the computation time of the SBP descriptors is sensibly higher than basic LBP^{riu2} or complemented LBP. Basically, this time is proportional to the size of the descriptor, and so will be still larger for higher order moments. However, it remains relatively low and faster than filter bank approaches such as Gabor filters [6], MR8 [7], Leung and Malik's [8] filters, or wavelets [10].

5.3. Results on the CUReT and UIUC datasets

Since the experiment settings as well as the observations obtained on those two datasets are similar, we present the results in the same section. In the experiments on the CUReT database, as in [2, 13], to get statistically significant experimental results, N training images were randomly chosen from each class while the remaining $92 - N$ images per class were used as the test set. Similarly, in experiments made on the UIUC database, N training images were randomly chosen from each class while the remaining $40 - N$ images per class were used as test set. The average classification rates on those databases over a hundred random splits with different parameters are

Table 2: Classification results on CURET dataset.

	(P,R)=(8,1)				(P,R)=(16,3)				(P,R)=(24,5)			
Number of training images N =	46	23	12	6	46	23	12	6	46	23	12	6
LTP [31]	85.13	79.25	72.25	63.09	92.66	87.30	80.22	70.50	91.81	85.78	77.88	67.77
$LBP^{riu2}/VAR_{P,R}$ [1]	93.87	88.76	81.59	71.03	94.20	89.12	81.64	71.81	91.87	85.58	77.13	66.04
CLBP_S/M [13]	93.52	88.67	81.95	72.30	94.45	90.40	84.17	75.39	93.63	89.14	82.47	73.26
CLBP_S/M/C [13]	95.59	91.35	84.92	74.80	95.86	92.13	86.15	77.04	94.74	90.33	83.82	74.46
BF + CLBP_S_M/C[51]	95.68	91.77	86.77	78.97	96.08	92.46	85.28	80.84	95.01	91.99	84.52	77.67
SBP_2	97.32	93.49	88.23	78.37	98.01	94.91	90.60	81.98	96.89	92.95	88.37	78.29
SBP_3	97.54	94.07	89.14	79.87	97.00	93.71	89.32	80.31	84.43	73.74	62.77	49.65
SBP_4	97.72	93.90	89.26	78.44	94.47	89.38	83.90	73.39	51.72	40.22	31.23	22.80
Number of training images N =	46				23				12			
MC_SBP_2	98.67				95.46				90.42			
MC_SBP_3	**98.73**				95.86				91.33			
MC_SBP_4	98.71				95.82				91.21			

reported respectively in Tables 2 and 3. Table 4 [3] compares the best results between different methods on these datasets.

As can be seen from these tables, the SBP descriptors significantly improve the best results of recent LBP-based methods on CURet and UIUC data sets (from 2% to 7%). For example, with $(P, R) = (16, 3)$, the best classification rates obtained by the CLBP with different numbers of training images (N = 46, 23, 12, or 6) on CUReT dataset are respectively 95.86%, 92.13%, 86.15%, and 77.04%. With the same configuration settings, the SBP descriptors reach classification rates of 98.01%, 94.91%, 90.60%, and 81.91%. The best result (98.73%) which is obtained with MC_SBP_3 improves the state-of-the-art performance on this dataset. For UIUC dataset, our best result (97.4%) outperforms LBP-based variants and it can be comparable to other state-of-the-art results.

[3] The nearest neighbour criterion is used for evaluating these methods. The results of [7], [11], [47], [48], [43], [49] and [50] are taken from [47].

Table 3: Classification results on UIUC dataset.

	(P,R)=(8,1)				(P,R)=(16,2)				(P,R)=(24,3)			
Number of training images N =	20	15	10	5	20	15	10	5	20	15	10	5
LBP^{riu2} [1]	54.65	52.94	47.14	39.72	61.32	56.42	51.16	42.67	64.05	60.05	54.25	44.59
CLBP_S/M [13]	81.80	78.55	74.8	64.84	87.87	85.07	80.59	71.64	89.18	87.42	81.95	72.53
CLBP_S/M/C [13]	87.64	85.70	82.65	75.05	91.04	89.42	86.29	78.57	91.19	89.21	85.95	78.05
CRLBP($\alpha = 1$) [20]	86.91	85.67	82.20	73.95	92.92	91.82	88.15	81.98	93.31	92.03	89.47	81.90
BF + CLBP_S/M [51]	91.21	87.15	83.19	75.65	93.45	90.99	87.97	80.19	93.78	91.64	88.12	80.23
SBP_2	91.31	89.56	85.97	78.19	95.52	94.34	91.87	85.69	96.55	95.40	93.07	87.11
SBP_3	93.43	91.79	88.47	80.77	96.35	95.27	92.91	86.96	97.09	96.00	93.78	88.21
SBP_4	91.55	89.74	86.36	77.75	94.94	93.46	90.75	83.37	95.99	94.77	92.16	85.53
Number of training images N =	20				15				10			
MC_SBP_2	97.04				96.03				94.00			
MC_SBP_3	**97.4**				96.35				94.11			
MC_SBP_4	95.64				94.66				92.16			

5.4. Results on KTH-TIPS and KTH-TIPS 2b datasets

Table 5 shows our results compared with different methods. Even using a small spatial support $(P, R) = (16, 2)$ at a single scale, our descriptor (SBP_3) reaches a classification accuracy of 98.1 %, which is higher than most of other methods. Considering multiple scales, the result can be further improved (98.29 %).

For KTH-TIPS 2b, we followed the training and testing scheme used in [44]. Experiments were done by training on one, two, or three samples (N_{train} on Table 6); testing is always performed on unseen samples only. For example, there exist six experiments where two samples are used as training and the two remaining samples are used as testing, and the classification rate reported is the average of those six obtained results.

Table 6 details the SBP results obtained at different orders and with different parameters. In addition, Figure 10 presents the best SBP results compared with the best results obtained by reference methods on KTH-TIPS 2b: LBP^{riu2} [1], VZ-MR8 [7], VZ-Joint [11], CLBC [16], CLBP [13]

and NI/RD/CI [55]. As it can be seen, the SBP framework significantly outperforms the competing algorithms, with an improvement which can reach 3.7%. Recently, Khan et al. [56] have significantly improved the state-of-the-art result (70.6%) on KTH-TIPS 2b dataset. However their method is based on a complex combination of many different texture descriptors: CLBP [13], WLD [25], BGP [57] , LPQ [58] and BSIF [59]. Furthermore, using multiscale approach, our descriptor (MC_SBP) still outperforms this result on KTH-TIPS 2b dataset, by reaching 71.59%. Only the combination between Improved Fisher Vector (IFV) and a deep learning technique (DeCAF) [45] outperforms this result. If they are considered separately, the results is not better than ours (IFV: 69.3, DeCAF: 70.7).

5.5. Experiment on DTD dataset

We have carried out an experiment on DTD dataset [45], compared with other LBP-based descriptors. We obtained a classification rate of 26 % with SBP using 1-NN classification. This result is better than other LBP-based methods using the same classification technique (see Table 7), but is significantly lower than the best methods on this dataset (66.7 % with a combination between IFV and DeCAF [45]). It must be said that this dataset shows an extreme variability in terms of appearance, scale and shape (see also Figure 9), that a simple, single-scale descriptor cannot represent adequately. While learning techniques used in [45] somehow compensate the descriptors for those variabilities, we associated our descriptors with the very simple Nearest Neighbour classifier (since the goal of the paper is to prove the

[4]Using IFV and DeCAF separately, the results are 69.3 and 70.7 respectively.

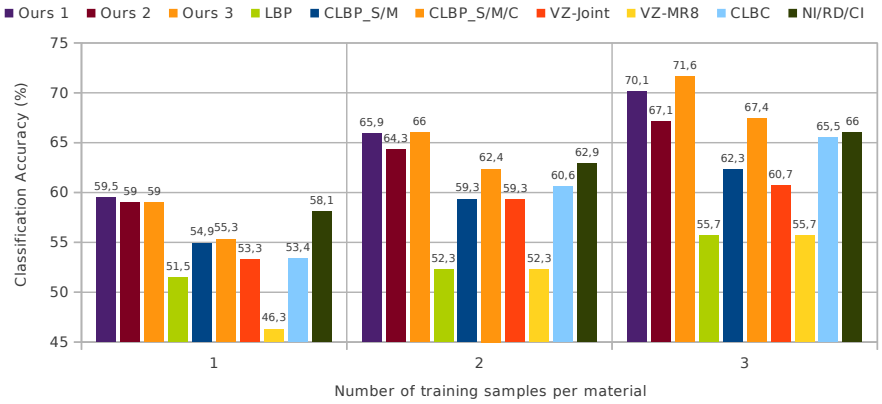


Figure 10: Comparing the best SBP descriptor on single scale (Ours 1), SBP descriptor using limited self-similarity $\text{LBP}^{riu2}_{8,1}$ (Ours 2) and multiscale complementary SBP descriptor (Ours 3) with state-of-the-art methods on on KTH-TIPS2b database. Ours 1: SBP_3 with $(P, R) = (24, 3)$, Ours 2: SBP_3 with $(P, R) = (8, 1)$ and Ours 3: MC_SBP_4 with $(P, R) = \{(8, 1), (8, 2), (8, 3), (8, 4), (8, 5)\}$.

strength of SBP descriptors only). Furthermore, the results on this dataset largely depend on machine learning techniques used for classification. According to the evaluation report [5] and [45], even with SVM classification, the results may significantly vary according to the used kernel. For example, using LBP method, the result is 7.35 % with linear SVM classifier, and 19.07 % with exp-chi2 kernel. If MR8 descriptor is used, the obtained result varies from 15.36 % with a linear SVM to 28.24 % with exp-chi2 kernel.

5.6. Global discussion

From the above experiments on five large datasets, we can derive a number of general conclusions:

[5] http://www.robots.ox.ac.uk/~vgg/data/dtd/eval.html

- The principle of SBPs is to extend the LBPs from the gray level to the local distribution, approximated by its first statistical moments. Being local filters, moment images can also be regarded as a filter bank preprocessing technique. The comparative evaluation shows that SBPs outperform recent state-of-the-art biologically-inspired filtering (BF) [51] for LBP-based operators (see BF+CLBP_S/M/C in Tables 2, 3).

- SBP_2 works well in all datasets and (P, R) configurations. It outperforms all competing LBP-based variants and is comparable to the state-of-the-art methods.

- SBP_3 and SBP_4 can perform better than SBP_2 when the spatial support of LBP is small ($R < 3$). The two preferred configurations are: $(P, R) = (8, 1)$ or $(P, R) = (16, 2)$. Moreover, SBP_3 proved more stable than SBP_4 in various datasets. Higher order moments should not be used this way when R is large. Indeed, higher order moment images have more local variations, which leads to a majority of non-uniform patterns. In our experiments, the configuration $(P, R) = (16, 2)$ proved a good trade-off for higher order SBP on the considered datasets.

- Using $\text{LBP}^{riu2}_{16,2}$ as self-similarity operator, SBP_3 outperforms the state-of-the-art results in many evaluations.

- Even with the smallest LBP spatial support ($\text{LBP}^{riu2}_{8,1}$), SBP_3 and SBP_4 can outperform or be comparable to the best results of other LBP-based variants and recent methods (see "Ours 2" in Figure 10). It implies that SBP descriptors can reduce considerably their computational cost with

respect to recent methods without decreasing the performance by using $\text{LBP}^{riu2}_{8,1}$ as self-similarity operator.

- The SBPs use only a single scale (P, R) of LBP operator. They do not need to consider multi-scale approach using different configurations of (P, R) like recent LBP-based variants (e.g. BRINT [36], NI/RD/CI [55]) to obtain state-of-the-art results, with the notable exception of the DTD dataset, whose variability is too high to be addressed by single-scale descriptors.

- The SBP framework provides efficient descriptors without using complementary magnitude information (CLBP_M), which is currently the essential factor for improving the performance in most state-of-the-art LBP-based variants, such as CLBP [13], CLBC [16], CRLBP [20], BRINT [36], disCLBP [40], SCLBP [41].

- Considering complementary multiscale approach, MC_SBP still improves the performance, notably in comparison with state-of-the-art results.

- Non-normalised moments, because they do not bring enough extra information with respect to the second order, do not improve the performance of SBP descriptors. Therefore only higher order SBP with normalised moments (SBP_3 and SBP_4) are recommended.

- Concerning the spatial support used to calculate moment images (structuring element), elements with two circles are recommended. Using smaller structuring elements make the statistics unsignificant, whereas

using bigger elements increase the feature extraction time without increasing the classification rate significantly. In practice, $\{(1,5),(2,8)\}$ is empirically chosen as structuring element.

6. Conclusion and perspectives

We have proposed a new collection of texture descriptors by extending the binary patterns from the pixel level to the local distribution level, applying rotation invariant self-similarity binary patterns to a set of moment images calculated using a structuring element as spatial support of the statistics. The interest of the proposed approach has been validated by a comparative evaluation on five large databases. It has been shown that the statistical binary patterns outperform the reference LBP variants, even using small LBP spatial support (e.g. $\text{LBP}^{riu2}_{8,1}$), and limiting the cardinality of the structuring element (e.g. around 15). Unlike most state-of-the-art LBP variants, the SBPs do not use complementary magnitude information (CLBP_M), but can be considered as a generalisation of the complementation approach, since it extends the self-similarity operator not only to local contrast information, but to higher order local variations.

The representation of local distribution using moment images, which is at the base of the SBPs, has a regularisation effect which makes the SBP more robust to small variations (e.g., noise) without reducing its discrimination power.

Finally, although our experiments have shown some benefits in using higher order (3 and 4) moments, two factors limit the interest of using them: (1) the computational cost due to the increased descriptor dimension, and

(2) their limited expresiveness due to the higher proportion of non uniform patterns. As a consequence, higher order moments are under-exploited in this study and more efficient ways to represent the local distribution should be investigated in the future.

References

[1] T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns, IEEE Trans. PAMI 24 (7) (2002) 971–987.

[2] S. Lazebnik, C. Schmid, J. Ponce, A sparse texture representation using local affine regions, IEEE Trans. PAMI 27 (8) (2005) 1265–1278.

[3] Z. Guo, Q. Li, L. Zhang, J. You, D. Zhang, W. Liu, Is Local Dominant Orientation Necessary for the Classification of Rotation Invariant Texture?, Neurocomputing 116 (2013) 182–191.

[4] X. Qi, R. Xiao, C. Li, Y. Qiao, J. Guo, X. Tang, Pairwise Rotation Invariant Co-Occurrence Local Binary Pattern, IEEE Trans. PAMI 36 (11) (2014) 2199–2213.

[5] X. Sun, J. Wang, M. F. She, L. Kong, Scale Invariant Texture Classification via Sparse Representation, Neurocomputing 122 (2013) 338–348.

[6] B. S. Manjunath, W.-Y. Ma, Texture Features for Browsing and Retrieval of Image Data, IEEE Trans. PAMI 18 (8) (1996) 837–842.

[7] M. Varma, A. Zisserman, A statistical approach to texture classification from single images, International Journal of Computer Vision 62 (1-2) (2005) 61–81.

[8] T. K. Leung, J. Malik, Representing and Recognizing the Visual Appearance of Materials using Three-dimensional Textons, International Journal of Computer Vision 43 (1) (2001) 29–44.

[9] J. Zhang, J. Liang, H. Zhao, Local Energy Pattern for Texture Classification Using Self-Adaptive Quantization Thresholds, IEEE Trans. on Image Processing 22 (1) (2013) 31–42.

[10] L. Sifre, S. Mallat, Rotation, Scaling and Deformation Invariant Scattering for Texture Discrimination, in: CVPR, 2013, pp. 1233–1240.

[11] M. Varma, A. Zisserman, A Statistical Approach to Material Classification Using Image Patch Exemplars, IEEE Trans. PAMI 31 (11) (2009) 2032–2047.

[12] S. Liao, M. W. K. Law, A. C. S. Chung, Dominant Local Binary Patterns for Texture Classification, IEEE Trans. Image Processing 18 (5) (2009) 1107–1118.

[13] Z. Guo, L. Zhang, D. Zhang, A Completed Modeling of Local Binary Pattern Operator for Texture Classification, IEEE Trans. Image Processing 19 (6) (2010) 1657–1663.

[14] N.-S. Vu, A. Caplier, Enhanced Patterns of Oriented Edge Magnitudes for Face Recognition and Image Matching, IEEE Trans. Image Processing 21 (2012) 1352–1365.

[15] N.-S. Vu, H. M. Dee, A. Caplier, Face Recognition using the POEM Descriptor, Pattern Recognition 45 (7) (2012) 2478–2488.

[16] Y. Zhao, D.-S. Huang, W. Jia, Completed local binary count for rotation invariant texture classification, IEEE Trans. Image Processing 21 (10) (2012) 4492–4497.

[17] L. Liu, L. Zhao, Y. Long, G. Kuang, P. Fieguth, Extended Local Binary Patterns for Texture Classification, Image Vision Comput. 30 (2012) 86–99.

[18] T. P. Nguyen, A. Manzanera, M. Garrigues, N. Vu, Spatial Motion Patterns: Action Models from Semi-Dense Trajectories, IJPRAI 28 (7).

[19] H. Jin, Q. Liu, H. Lu, X. Tong, Face Detection using Improved LBP under Bayesian Framework, in: ICIG, 2004, pp. 306–309.

[20] Y. Zhao, W. Jia, R.-X. Hu, H. Min, Completed Robust Local Binary Pattern for Texture Classification, Neurocomputing 106 (2013) 68–76.

[21] S. Liao, X. Zhu, Z. Lei, L. Zhang, S. Li, Learning Multi-scale Block Local Binary Patterns for Face Recognition, in: ICB, Vol. 4642 of LNCS, 2007, pp. 828–837.

[22] M. Heikkila, M. Pietikainen, C. Schmid, Description of interest regions with local binary patterns, Pattern Recognition 42 (2009) 452–436.

[23] G. Zhao, T. Ahonen, J. Matas, M. Pietikäinen, Rotation-Invariant Image and Video Description With Local Binary Pattern Features, IEEE Trans. Image Processing 21 (4) (2012) 1465–1477.

[24] R. Maani, S. Kalra, Y.-H. Yang, Rotation Invariant Local Frequency Descriptors for Texture Classification, IEEE Trans. on Image Processing 22 (6) (2013) 2409–2419.

[25] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikäinen, X. Chen, W. Gao, WLD: A Robust Local Image Descriptor, IEEE Trans. PAMI 32 (9) (2010) 1705–1720.

[26] W. Zhang, S. Shan, W. Gao, X. Chen, H. Zhang, Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-Statistical Model for Face Representation and Recognition, in: ICCV, 2005, pp. 786–791.

[27] S. Liao, A. C. S. Chung, Face Recognition by using Elongated Local Binary Patterns with Average Maximum Distance Gradient Magnitude, in: Y. Yagi, S. B. Kang, I.-S. Kweon, H. Zha (Eds.), ACCV, Vol. 4844 of LNCS, 2007, pp. 672–679.

[28] L. Wolf, T. Hassner, Y. Taigman, Descriptor based Methods in the Wild, in: Real-Life Images Workshop, ECCV, 2008.

[29] T. Mäenpää, M. Pietikäinen, Multi-scale Binary Patterns for Texture Analysis, in: SCIA, Vol. 2749 of LNCS, 2003, pp. 885–892.

[30] G. Zhao, M. Pietikainen, X. Chen, Combining LBP Difference and Feature Correlation for Texture Description, IEEE Trans. Image Processing (2014) 2557–2568.

[31] X. Tan, B. Triggs, Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions, IEEE Trans. Image Processing 19 (6) (2010) 1635–1650.

[32] L. Nanni, A. Lumini, S. Brahnam, Local Binary Patterns Variants as Texture Descriptors for Medical Image Analysis, Artificial Intelligence in Medicine 49 (2) (2010) 117–125.

[33] M. A. Akhloufi, A. Bendada, Locally Adaptive Texture Features for Multispectral Face Recognition, in: SMC, 2010, pp. 3308–3314.

[34] S. Liao, G. Zhao, V. Kellokumpu, M. Pietikäinen, S. Z. Li, Modeling Pixel Process with Scale Invariant Local Patterns for Background Subtraction in Complex Scenes, in: CVPR, 2010, pp. 1301–1306.

[35] A. Hafiane, G. Seetharaman, B. Zavidovique, Median Binary Pattern for Textures Classification, in: ICIAR, 2007, pp. 387–398.

[36] L. Liu, Y. Long, P. W. Fieguth, S. Lao, G. Zhao, BRINT: Binary Rotation Invariant and Noise Tolerant Texture Classification, IEEE Trans. on Image Processing 23 (7) (2014) 3071–3084.

[37] J. Ren, X. Jiang, J. Yuan, Noise-Resistant Local Binary Pattern With an Embedded Error-Correction Mechanism, IEEE Trans. Image Processing 22 (10) (2013) 4049–4060.

[38] D. K. Iakovidis, E. G. Keramidas, D. Maroulis, Fuzzy Local Binary Patterns for Ultrasound Texture Characterization, in: ICIAR, 2008, pp. 750–759.

[39] T. P. Nguyen, A. Manzanera, Incorporating Two First Order Moments into LBP-Based Operator for Texture Categorization, in: C. V. Jawahar, S. Shan (Eds.), ACCV Workshops, Vol. 9008 of LNCS, 2014, pp. 527–540.

[40] Y. Guo, G. Zhao, M. Pietikäinen, Discriminative Features for Texture Description, Pattern Recognition 45 (10) (2012) 3834–3843.

[41] V. D. Nguyen, D. D. Nguyen, T. T. Nguyen, V. Q. Dinh, J. W. Jeon, Support Local Pattern and its Application to Disparity Improvement and Texture Classification, IEEE Trans. Circuits Syst. Video Techn. 24 (2) (2014) 263–276.

[42] K. J. Dana, B. van Ginneken, S. K. Nayar, J. J. Koenderink, Reflectance and Texture of Real-World Surfaces, ACM Trans. Graph. 18 (1999) 1–34.

[43] E. Hayman, B. Caputo, M. Fritz, J. Eklundh, On the Significance of Real-World Conditions for Material Classification, in: T. Pajdla, J. Matas (Eds.), ECCV, Vol. 3024 of LNCS, 2004, pp. 253–266.

[44] B. Caputo, E. Hayman, M. Fritz, J.-O. Eklundh, Classifying Materials in the Real World, Image Vision Comput. 28 (1) (2010) 150–163.

[45] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, A. Vedaldi, Describing Textures in the Wild, in: CVPR, 2014, pp. 3606–3613.

[46] T. Ojala, T. Mäenpää, M. Pietikäinen, J. Viertola, J. Kyllönen, S. Huovinen, Outex - New Framework for Empirical Evaluation of Texture Analysis Algorithms, in: ICPR, 2002, pp. 701–706.

[47] L. Liu, P. W. Fieguth, D. A. Clausi, G. Kuang, Sorted Random Projections for Robust Rotation-invariant Texture Classification, Pattern Recognition 45 (6) (2012) 2405–2418.

[48] J. Zhang, M. Marszalek, S. Lazebnik, C. Schmid, Local Features and Kernels for Classification of Texture and Object Categories: A Comprehensive Study, Int. J. Comput. Vision 73 (2) (2007) 213–238.

[49] Y. Xu, H. Ji, C. Fermüller, Viewpoint Invariant Texture Description Using Fractal Analysis, Int. J. Comput. Vision 83 (1) (2009) 85–100.

[50] M. Crosier, L. D. Griffin, Using Basic Image Features for Texture Classification, Int. J. Comput. Vision 88 (3) (2010) 447–460.

[51] N. Vu, T. P. Nguyen, C. Garcia, Improving Texture Categorization with Biologically-inspired Filtering, Image Vision Comput. 32 (6-7) (2014) 424–436.

[52] L. Liu, P. W. Fieguth, Texture Classification from Random Features, IEEE Trans. PAMI 34 (3) (2012) 574–586.

[53] F. M. Khellah, Texture Classification Using Dominant Neighborhood Structure, IEEE Trans. on Image Processing 20 (11) (2011) 3270–3279.

[54] L. Zhang, L. Zhang, Z. Guo, D. Zhang, Monogenic-LBP: a New Approach for Rotation Invariant Texture Classification, in: ICIP, 2010, pp. 2677–2680.

[55] L. Liu, L. Zhao, Y. Long, G. Kuang, P. W. Fieguth, Extended Local Binary Patterns for Texture Classification, Image Vision Comput. 30 (2) (2012) 86–99.

[56] F. S. Khan, R. M. Anwer, J. van de Weijer, M. Felsberg, J. Laaksonen,

Compact Color-texture Description for Texture Classification, Pattern Recognition Letters 51 (2015) 16–22.

[57] L. Zhang, Z. Zhou, H. Li, Binary Gabor Pattern: An Efficient and Robust Descriptor for Texture Classification, in: ICIP, 2012, pp. 81–84.

[58] E. Rahtu, J. Heikkilä, V. Ojansivu, T. Ahonen, Local Phase Quantization for Blur-insensitive Image Analysis, Image Vision Comput. 30 (8) (2012) 501–512.

[59] J. Kannala, E. Rahtu, BSIF: Binarized Statistical Image Features, in: ICPR, 2012, pp. 1363–1366.

[60] R. Timofte, L. J. V. Gool, A Training-free Classification Framework for Textures, Writers, and Materials, in: BMVC, 2012, pp. 1–12.

Biography



Thanh Phuong Nguyen received the Ph.D. degree in computer science from University of Nancy 1, Nancy, in 2010. He has been a Teaching Assistant with University of Nancy 1 and University of Nancy 2, Nancy, from 2009 to 2011. He was a Post-Doctoral Researcher with the Center for Mathematical Morphology, Mines ParisTech, Paris, France, in 2012 and then with the Robotics and Computer Vision Laboratory, National School of Advanced Technologies, Palaiseau, France from 2012 to 2015. Since 2015, he has been an Ass. Prof. at Southern University of Toulon-Var, France. His research

interests include image analysis, visual descriptor, action recognition, texture analysis and shape representation.



Ngoc-Son Vu is currently an Ass. Prof. at ENSEA, Cergy-Pontoise, France and a member of the ETIS-ENSEA/UCP/CNRS laboratory. He worked as a researcher at INSA de Lyon, France (in 2013) and at the Vesalis company, Clermont-Ferrand, France (in 2012). He was also a postdoctoral researcher at Grenoble Institute of Technology, France (in 2011) from which he received his Ph.D. degree in image processing and computer vision in 2010. His research interests include image processing, pattern recognition, texture, human body and face analysis.



Antoine Manzanera received his BS degree in Mathematics in 1991 and his M.S. degree in Theoretical Computer Science in 1993 from Lyon I University, his Ph.D. from Tlcom ParisTech in 2000, and Habilitation degree (HDR) from Paris VI University in 2012. He has been a high school teacher in Paraguay (from 19941996) and a R&D Engineer (CIFRE grant) at Arospatiale-Missiles (from 19972000). Since 2001, he has been an Associate Professor at ENSTA-ParisTech in the Robotics & Computer Vision group from the U2IS laboratory. His research domains are image models and video processing algorithms, with a particular interest in motion modeling and analysis for mobile embedded systems.

Method	CUReT (N=46)	UIUC (N=20)
Ours (MC_SBP)	98.73	97.4
$dis(S+M)^{ri}_{N,R}$ [40]	98.3	
VZ_MR8 [7]	97.43	
VZ_Joint [11]	98.03	97.83
SRP [47]	98.52	96.27
Lazebnik [2]	72.5	96.03
Multiscale BIF [50]	98.6	98.8
Hayman [43]	98.46	92.0
RP [52]	98.52	
BRINT2_S_M (MS9) [36]	97.86	
DNS+$LBP_{24,3}$ [53]	94.52	
Zhang et al. [48]	95.3	98.7
Xu et al. [49]		92.74
Joint_Sort [3]	96.93	92.73
PRICoLBP [4]	98.4	
Mono_LBP [54]	98.27	95.44

Table 4: Comparison between the best results of different methods on CUReT and UIUC dataset.

Table 5: Classification rates obtained on KTH-TIPS dataset.

Method	SBP_2	SBP_3	SBP_4	M_SBP_2	[9]	[4]	[47]
Result	97.18	98.10	97.63	98.29	96.41	98.4	97.71

Table 6: Classification rates obtained on the KTH-TIPS 2b database.

(P,R)	(8,1)			(16,2)			(24,3)		
N_{train}	1	2	3	1	2	3	1	2	3
LBP [1]	48.1	54.2	52.6	50.5	55.8	59.1	49.9	54.6	57.8
NI/RD/CI [55]	56.6	61.9	64.8	57.7	62.5	65.1	52.4	57.5	61.7
CLBP_S/M [13]	53.43	58.26	60.42	54.92	59.33	62.29	54.30	58.96	61.91
CLBP_S/M/C [13]	53.28	60.89	64.90	55.06	62.03	67.11	55.32	62.37	67.42
CLBC_S/M [16]	50.29	55.42	57.87	50.84	55.20	57.58	50.16	54.93	58.21
CLBC_S/M/C [16]	52.04	59.60	63.68	53.37	60.61	65.49	52.53	59.46	64.14
(P,R)	{(8,1),(16,2)}			{(16,2),(24,3)}			{(8,1),(16,2),(24,3)}		
NI/RD/CI [55]	58.1	62.9	66.0	55.9	61.0	64.2	56.7	61.7	65.0
$N_{train}=$	1			2			3		
VZ-Joint	53.3			59.3			60.7		
VZ-MR8	46.3			52.3			55.7		
CLBP + WLD + BGP + LPQ + BSIF [56]							70.6		
Timofte and Van Gool [60]							66.3		
Cimpoi et al. [45] DeCAF + IFV							76.2 [4]		
Our descriptors									
	(P,R)=(8,1)			(P,R)=(16,2)			(P,R)=(24,3)		
N_{train}	1	2	3	1	2	3	1	2	3
SBP_2	57.68	63.24	66.03	59.03	64.91	68.60	59.04	65.20	68.81
SBP_3	58.98	64.30	67.07	60.08	65.63	69.61	59.46	65.86	70.08
SBP_4	55.16	60.97	65.19	57.07	63.26	66.96	56.74	62.75	67.07
N_{train}	1			2			3		
MC_SBP_2	57.74			64.04			68.54		
MC_SBP_3	58.35			65.12			70.43		
MC_SBP_4	59.03			65.99			**71.59**		

Table 7: Classification rate on the DTD dataset with different LBP-based methods.

Method	$LBP^{riu2}_{(24,3)}$	$CLBP^{riu2}_{(24,3)}S/M/C$	$SBP_2(24,3)$	$CSBP_2(24,3)$
Classification rate	14.51	20.49	26.18	26.38