



Extended local binary patterns for texture classification [☆]

Li Liu ^{a,*}, Lingjun Zhao ^a, Yunli Long ^a, Gangyao Kuang ^a, Paul Fieguth ^b

^a School of Electronic Science and Engineering, National University of Defense Technology, 47 Yanwachi, Changsha, Hunan, 410073, China

^b Department of Systems Design Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

ARTICLE INFO

Article history:

Received 19 April 2011

Received in revised form 20 October 2011

Accepted 4 January 2012

Keywords:

Texture classification
Local binary pattern (LBP)
Bag-of-words (BoW)
Rotation invariance

ABSTRACT

This paper presents a novel approach for texture classification, generalizing the well-known local binary pattern (LBP) approach. In the proposed approach, two different and complementary types of features (pixel intensities and differences) are extracted from local patches. The intensity-based features consider the intensity of the central pixel (CI) and those of its neighbors (NI); while for the difference-based feature, two components are computed: the radial-difference (RD) and the angular-difference (AD). Inspired by the LBP approach, two intensity-based descriptors CI-LBP and NI-LBP, and two difference-based descriptors RD-LBP and AD-LBP are developed. All four descriptors are in the same form as conventional LBP codes, so they can be readily combined to form joint histograms to represent textured images. The proposed approach is computationally very simple: it is totally training-free, there is no need to learn a texton dictionary, and no tuning of parameters. We have conducted extensive experiments on three challenging texture databases (Outex, CURET and KTHTIPS2b). Outex results show significant improvements over the classical LBP approach, which clearly demonstrates the great power of the joint distributions of these proposed descriptors for gray-scale and rotation invariant texture classification. The proposed method produces the best classification results on KTHTIPS2b, and results comparable to the state-of-the-art on CURET.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Texture classification is a fundamental issue in computer vision and image processing, playing a significant role in a wide range of applications that include medical image analysis, remote sensing, object recognition, document analysis, environment modeling, content-based image retrieval etc. [1]. For four decades, texture analysis has been an area of intense research, however analyzing real world textures has proven to be surprisingly difficult, in many cases caused by natural texture inhomogeneity of varying illumination, scale changes, and variability in surface shape.

Recently, the orderless Bag-of-Words (BoW) [5, 2, 3, 8] approach, representing texture images statistically as histograms over a discrete texton dictionary, has proven extremely popular and successful in texture classification tasks. Robust and discriminative *local* texture descriptors and *global* statistical histogram characterization have supplied complementary components toward the BoW feature extraction of texture images. The former attempts to extract a collection of powerful and distinctive appearance descriptors from local patches; while the latter first utilizes the fact that texture images contain self-repeating patterns by vector-quantifying (typically by *k*-means) the local feature

vectors to form a texton dictionary, and then represent texture images statistically as compact histograms over the learned texton dictionary.

In this simple and efficient BoW framework, it is generally agreed that the local descriptors play a much more important role, and have therefore received considerable attention [2–5, 8, 9, 6]. The approaches can be grouped into sparse and dense types, with the sparse approach using appearance descriptors at a sparse set of detected interest points. Noticeable sparse descriptors include SPIN, SIFT and RIFT [8, 10]. In contrast, dense approaches use appearance descriptors pixel by pixel [2–5, 9]. The sparse approach largely relies on the sparse output of local interest region detectors, which might miss important texture primitives and fail to provide enough regions for a robust statistical characterization of the texture.

Among the most popular dense descriptors are the various filter banks, such as Gabor filters [11], the filter bank of Schmid [5], the filter bank of Leung and Malik [5], the MR8 [2], the filter bank of Crosier [9] and many others [12]. The design of a filter bank is nontrivial and likely to be application dependent. Although enormous efforts have been carried out along this direction, the supremacy of filter bank-based descriptors for texture analysis has been challenged by several authors [4, 3, 7] who have demonstrated that using the intensities or differences in a local small patch directly can produce superior or comparable classification performance to filter banks with large spatial support. In [7], the authors propose sparse modeling of local texture patches, however the sparse texton learning and sparse coding process is computationally expensive. Two particularly important works along these lines are the VZ-joint classifier [3] and the LBP method [4]. The simple, elegant and

[☆] This paper has been recommended for acceptance by Matti Pietikainen.

* Corresponding author. Tel.: +86 731 84573479 807; fax: +86 731 84518730.

E-mail addresses: dreamliu2010@gmail.com (L. Liu), lingjun.zhao@gmail.com (L. Zhao), feiyunli@hotmail.com (Y. Long), kuangyeats@vip.sina.com (G. Kuang), pfieguth@uwaterloo.ca (P. Fieguth).

efficient local texture descriptor LBP may be the preferable choice over VZ-Joint classifier since LBP uses a pre-defined texton dictionary and does not need to use nearest neighbor to obtain the texton labels, a time consuming step.

Due to its impressive computational efficiency and good texture discriminative property, the dense LBP descriptor [4] has gained considerable attention since its publication [13], and has already been used in many other applications, including visual inspection, image retrieval, dynamic texture recognition, remote sensing, biomedical image analysis, face image analysis, motion analysis, environment modeling, and outdoor scene analysis [14–16, 18, 19, 34].¹ Despite the great success of LBP in computer vision and pattern recognition, the conventional LBP operator comes with disadvantages and limitations:

1. The LBP operator produces long histograms which are sensitive to image rotation.
2. The LBP has small spatial support; in its basic form, the LBP operator cannot properly detect large-scale textural structures.
3. LBP loses local textural information, since only the signs of differences of neighboring pixels are utilized.
4. LBP is very sensitive to noise. The slightest fluctuation above or below the value of the central pixel is treated as equivalent to a major contrast between the central pixel and its surroundings.

On the basis of the above issues, researchers have proposed a variety of LBP variants. In terms of locality, the authors in [20] propose to extract global features from Gabor filter responses as a complementary descriptor. In order to recover the loss of information created by computing the LBP value, the local image contrast has been introduced by Ojala et al. [4] as a complementary measure, and better performance has been reported therein. Moreover, Guo et al. [21] propose to include the information contained in the magnitudes of local differences as complementary to the signs used by LBP, and claim better performance.

Regarding LBP robustness, especially to noise, the influential work by Ojala et al. [4] extends basic LBP to a multiresolution context, and rotation invariant patterns are introduced and successfully used in reducing the dimension of the LBP histogram and enhancing robustness and speed. Ahonen et al. introduce soft histograms [28], and Tan and Triggs [29] introduced local ternary patterns (LTP), using tertiary numbers instead of binary. Noting that uniform LBPs are not necessary to occupy the major pattern proportions, Liao et al. [20] proposed to use dominant LBP (DLBP) which considers the most frequently occurred patterns in a texture image.

Very recently, Heikkilä et al. [22] exploit circular symmetric LBP (CS-LBP) for local interest region description, and Chen et al. present a WLD descriptor by including orientation information as a robust descriptor [23].

The LBP approach is based on the assumption that the local differences of the central pixel and its neighbors are independent of the central pixel itself. However, in practice an exact independence is not warranted: the superiority of both VZ-Joint and VZ-MRF classifiers over LBP clearly demonstrates the benefits of explicitly including the information contained in the central pixel [3].

The fundamental question being raised here is whether explicitly modeling the joint distribution of the central pixel and its neighbors is an advantage or not, and how to effectively include the missing between-scale information so that better texture classification can be achieved? Motivated by the work of Varma and Zisserman [3] and the LBP approach studied by Ojala et al. [4], in this paper we propose a simple, yet very powerful and novel local texture descriptor to generalize the conventional LBP approach. In the proposed approach, two different but complementary types of features in a local patch, the pixel intensities and the pixel differences, are utilized by using a

common concept, the LBP coding strategy. The pixel intensities are divided into two components: the intensity of the central pixel and the intensities of its neighboring pixels. For pixel differences, we study radial and angular differences.

All four descriptors (two intensity based, two difference based) are in the same form as the conventional LBP codes, thus they can be readily combined to form a joint histogram. The fusing of these descriptors will be shown to lead to significantly improved classification results on the experimental protocols designed for verifying the performance of the LBP approach in [4]. The key to our proposed approach is that it employs the advantages of VZ-Joint/VZ-MRF in its strong performance from having a joint distribution, and those of LBP in computational efficiency.

The paper is organized as follows: we start with a brief review of the classical LBP approach in Section 2, followed by details of the derivation of the proposed descriptors and the classification scheme. In Section 3, we verify the proposed approach with extensive experiments on popular texture datasets and comparisons with various state-of-the-art texture classification techniques. Section 4 provides concluding remarks and possible extensions of the proposed method. A short, preliminary version of this work appeared in [26].

2. Proposed descriptors

This section begins by reviewing conventional LBP, followed by the new descriptors designed to address the limitations of LBP. Finally, the multiresolution analysis and classification scheme of this work is presented.

2.1. A brief review of LBP

The LBP method, first proposed by Ojala et al. [25, 4], encodes the pixel-wise information in textured images. Images are probed locally by sampling grayscale values at a central point $x_{0,0}$ and p points $x_{r,0}, \dots, x_{r,p-1}$ spaced equidistantly around a circle of radius r (the choice of which acts as a surrogate for controlling the scale of description), as shown in Fig. 1. In LBP, a “local pattern” operator describes the relationships between a pixel and its neighborhood pixels; all neighbors that have values higher than or equal to the value of the central pixel are given a value of 1, and all those lower a value of 0. The binary values associated with the neighbors are then read sequentially, clockwise, to form a binary number which may be used to characterize the local texture. Formally,

$$LBP_{p,r} = \sum_{n=0}^{p-1} s(x_{r,n} - x_{0,0}) 2^n, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1)$$

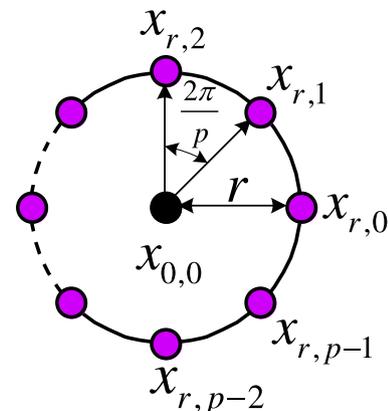


Fig. 1. A central pixel $x_{0,0}$ and its p circularly and evenly spaced neighbors $\{x_{r,i}\}_{i=0}^{p-1}$ on radius r .

¹ A bibliography of LBP-related research can be found at http://www.cse.oulu.fi/MVG/LBP_Bibliography/.

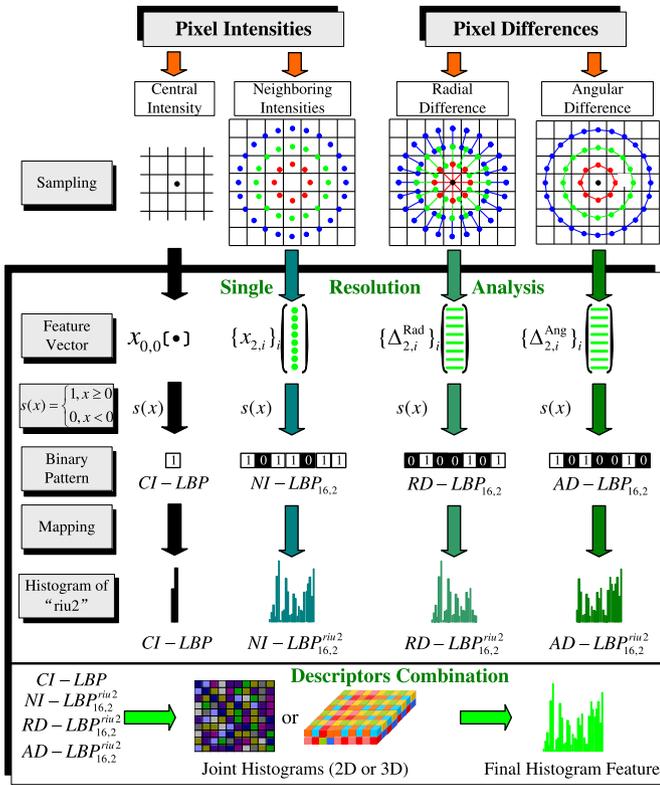


Fig. 2. Overview of the proposed approach.

Relative to the origin at $(0,0)$, the coordinates of the neighbors are given by $r \sin(2\pi m/p), r \cos(2\pi m/p)$. The gray values of neighbors which do not fall exactly in the center of pixels are estimated by interpolation.

Given an $N \times M$ image \mathbf{I} , let $LBP_{p,r}(i,j)$ be the identified LBP pattern of each pixel (i,j) , then the whole textured image is represented by a histogram vector h of length K :

$$h(k) = \sum_{i=1}^N \sum_{j=1}^M \delta(LBP_{p,r}(i,j) - k) \quad (2)$$

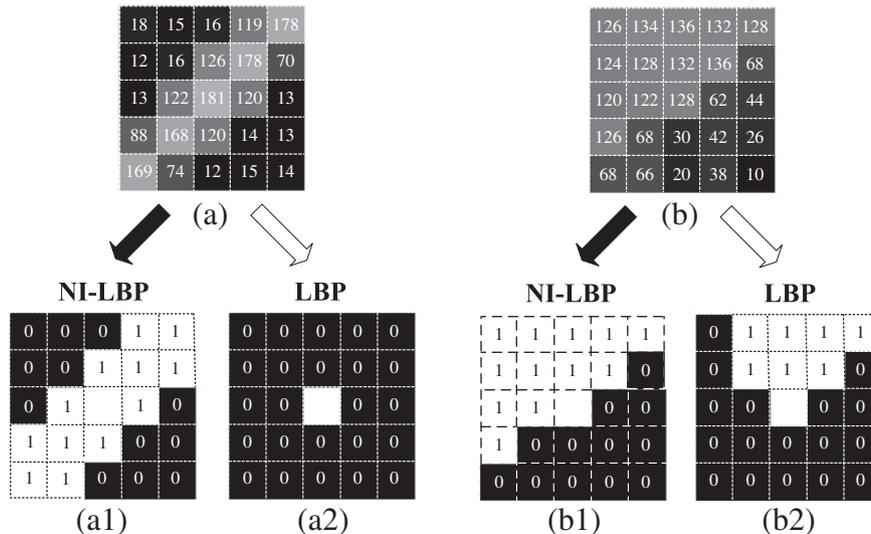


Fig. 3. The proposed NI-LBP preserves weak edge patterns. Two 5×5 example image patches are shown in (a,b). (a1) NI-LBP patterns of (a) at two resolutions with thresholds of 108 and 52.4 respectively. (b1) NI-LBP patterns of (b) at two resolutions with thresholds of 90 and 85.4 respectively. (a2,b2) are the patterns given by LBP.

where $0 \leq k \leq K - 1$, $K = 2^p$ is the number of LBP codes, and $\delta(\bullet)$ is the Dirac delta function. This formulation has several attractive properties that favor its usage: gray-scale invariance, computational speed, few parameters, satisfactory discriminant power, and rotation invariance achieved by simple cyclic shifts.

On the other hand, the basic LBP operator produces rather long histograms (2^p distinct values), and it can become intractable to estimate h due to the overwhelming dimensionality of h with large p . Moreover, it is easy to realize that due to the way LBP numbers are created, they are very sensitive to noise: the slightest fluctuation above or below the value of the central pixel is treated the same way as a major contrast between the central pixel and its surroundings. One way to avoid noisy patterns is to simply ignore them: a noisy pattern due to its randomness will create neighbors that fluctuate above and below the value of the central pixel, with 0s and the 1s frequently succeeding each other. Therefore, one improvement suggested by Ojala et al. [4] is to consider only the so-called “uniform” patterns by proposing the $LBP_{p,r}^{riu2}$ operator, merging nonuniform patterns directly into one pattern. The success of the $LBP_{p,r}^{riu2}$ operator also comes from that fact that the “uniform” patterns appear to be fundamental properties of local image textures [4], and represent some prominent and salient local texture structures. The $LBP_{p,r}^{riu2}$ operator is formally defined as

$$LBP_{p,r}^{riu2} = \begin{cases} \sum_{n=0}^{p-1} s(x_{r,n} - x_{0,0}), & \text{if } U(LBP_{p,r}) \leq 2 \\ p + 1, & \text{otherwise} \end{cases} \quad (3)$$

where

$$U(LBP_{p,r}) = \sum_{n=0}^{p-1} |s(x_{r,n} - x_{0,0}) - s(x_{r, \text{mod}(n+1,p)} - x_{0,0})| \quad (4)$$

where superscript *riu2* denotes the rotation invariant “uniform” patterns that have U values at most 2. Therefore, mapping from $LBP_{p,r}$ to $LBP_{p,r}^{riu2}$ results in only $p + 1$ distinct groups of patterns, leading to a much shorter histogram representation for the whole image.

It is obvious that LBP oversimplifies local structure and loses textural information. Therefore, Ojala et al. [4] made a further important correction by including the local contrast of each pattern and proposing a complementary local descriptor called $VAR_{p,r}$. Using the 2D joint histogram of $LBP_{p,r}^{riu2}$ and $VAR_{p,r}$, denoted as $LBP_{p,r}^{riu2}/VAR_{p,r}$ is demonstrated in [4].

In conventional LBP the central pixel is discarded (despite the implicit use of the intensity of the central pixel as the threshold to achieve local gray-scale invariance), and only the joint distribution of the neighborhood around each pixel is considered. However, in their recent extensive texture study, Zhang et al. [8] suggested that it is vital to use a combination of several detectors and descriptors. Motivated by the work of Lazebnik et al. [10] and Zhang et al. [8], in this paper we seek to propose a method which possesses the strengths of combining complementary local features, with those of LBP in computational efficiency and smaller support regions.

2.2. Intensity-based descriptors

The brightness level at a point in an image is highly dependent on the brightness levels of its neighboring points unless the image is simply random noise [24]. In MRF modeling [24], the probability of the central pixel depends only on its neighborhood as

$$p(\mathbf{I}(x_c) | \mathbf{I}(x), \forall x \neq x_c) = p(\mathbf{I}(x_c) | \mathbf{I}(x), \forall x \in \mathcal{N}(x_c)) \tag{5}$$

where x_c is a site in the 2D integer lattice on which the image \mathbf{I} has been defined and $\mathcal{N}(x_c)$ is the neighborhood of that site. The center pixel also has discriminant information, however its distribution is conditioned on its neighbors alone.

Inspired by such MRF models, and related to the ideas explored by Varma and Zisserman [3], we propose to use only local neighborhood

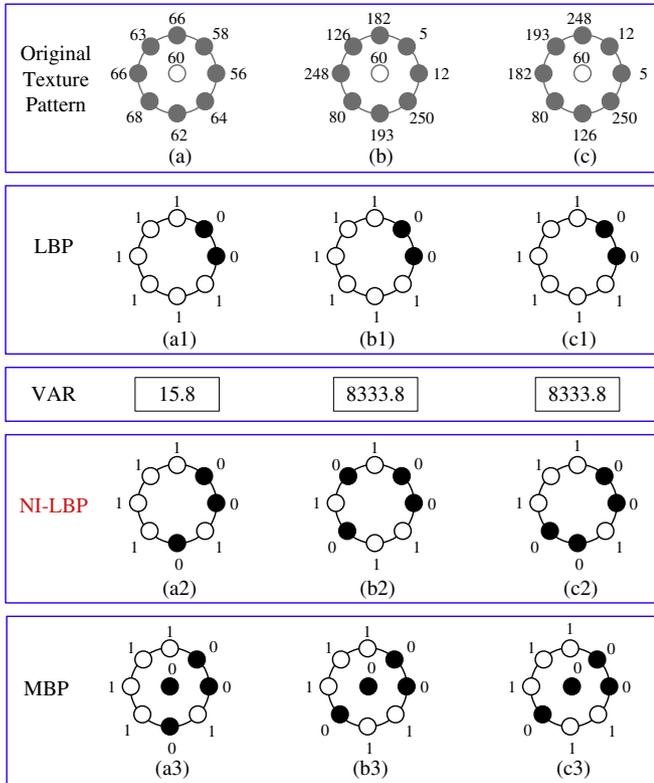


Fig. 4. Three different original texture patterns (a, b, c) and their corresponding LBPs (a1, b1, c1), NI-LBPs (a2, b2, c2), MBPs (a3, b3, c3) and VAR values. All three LBP patterns (a1, b1, c1) are the same. Patterns in (a) and (b) would be considered as the same pattern type by LBP, though corresponding textural surfaces might be quite different from each other. By incorporating LBP with local variance information, patterns in (a) and (b) could be distinguished, while patterns in (b) and (c) would still be considered as the same pattern type because of the same variance. But they are different in configuration, which is not due to the rotation but underlying textural properties. In terms of MBP, MBP can distinguish (a) and (b). However, MBP cannot distinguish (b) and (c). In contrast, all three NI-LBPs are different. Therefore, all three patterns can be distinguished by our proposed NI-LBP.

Table 1

Comparison of the detail mean classification accuracy for NI-LBP, LBP and MBP on test suite Outex_TC_00000. Results are obtained as the average of the 100 test groups. The patch size used is 3×3 . The 1NN classifier is used. The distance measure is χ^2 . Each image is normalized to have zero mean and unit standard deviation. “NI-LBP (512)” encodes the binary value of the center pixel, similar to “MBP (512)”. “MBP (256)” excludes the binary value of the center pixel. Since LBP uses the value of the center pixel as the threshold, therefore it is unnecessary to include the center pixel in this case. The numbers in the brackets denote the number of bins of the histogram.

Class	NI-LBP (512)	NI-LBP (256)	LBP	MBP (512)	MBP (256)
canvas001	100.0%	100.0%	100.0%	100.0%	100.0%
canvas002	100.0%	100.0%	100.0%	100.0%	100.0%
canvas003	100.0%	100.0%	100.0%	100.0%	100.0%
canvas005	100.0%	100.0%	100.0%	100.0%	100.0%
canvas006	100.0%	100.0%	100.0%	100.0%	100.0%
canvas009	100.0%	100.0%	100.0%	100.0%	100.0%
canvas011	100.0%	100.0%	100.0%	100.0%	100.0%
canvas021	100.0%	100.0%	100.0%	100.0%	100.0%
canvas022	100.0%	100.0%	100.0%	100.0%	100.0%
canvas023	99.6%	99.4%	99.6%	99.8%	99.8%
canvas025	100.0%	100.0%	100.0%	100.0%	100.0%
canvas026	100.0%	100.0%	100.0%	100.0%	100.0%
canvas031	100.0%	100.0%	100.0%	100.0%	100.0%
canvas032	100.0%	100.0%	100.0%	100.0%	100.0%
canvas033	95.5%	96.0%	92.0%	94.4%	92.5%
canvas035	100.0%	100.0%	100.0%	100.0%	100.0%
canvas038	95.5%	98.2%	100.0%	99.7%	99.6%
canvas039	99.8%	99.5%	100.0%	99.6%	99.8%
tile005	100.0%	100.0%	100.0%	100.0%	100.0%
tile006	100.0%	99.6%	100.0%	99.8%	99.7%
carpet002	100.0%	100.0%	100.0%	100.0%	100.0%
carpet004	100.0%	100.0%	100.0%	100.0%	100.0%
carpet005	100.0%	99.6%	100.0%	99.4%	96.5%
carpet009	99.9%	99.9%	99.3%	95.2%	94.4%
Mean	99.76%	99.68%	99.62%	99.50%	99.26%

distributions in our NI-LBP descriptor. We explicitly model the joint distribution of the central pixels and its neighbors, in order to test how significant this conditional probability distribution is for classification.

Next, inspired by the coding strategy of LBP, we define the following NI-LBP descriptor (see also Fig. 2):

$$NI-LBP_{p,r} = \sum_{n=0}^{p-1} s(x_{r,n} - \mu) 2^n, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{6}$$

where $\mu = \frac{1}{2} \sum_{n=0}^{p-1} x_{r,n}$. Similar to $LBP_{p,r}^{riu2}$, the rotation invariant version of NI-LBP, denoted by $NI-LBP_{p,r}^{riu2}$, can also be defined to achieve rotation invariant classification.

Regarding the selection of the threshold μ , although it was motivated by intuition and experimental studies, it is also selected in order to preserve LBP characteristics and to increase robustness. Hafiane et al. [17] proposed Median Binary Pattern (MBP) which seeks to derive the localized binary pattern by thresholding the pixels against their median value over a 3×3 neighborhood. In MBP, the central pixel is also included in this filtering process, resulting 2^9 binary patterns.

NI-LBP, LBP and MBP differ in the selection of thresholding value. The capability of encoding image configuration and pixelwise relationships might be different as they use different thresholds. For illustration purpose, Fig. 4 gives three different example local texture patterns. The patterns shown in Fig. 4(a) and (b) would be classified into the same class. But the textural surfaces they represent are quite different from each other, which means they probably belong to different classes. While the other three descriptors NI-LBP, MBP and VAR can all tell the difference between (a) and (b). This is why Ojala et al. use the combination of LBP and VAR. However, the joint histogram of LBP and VAR cannot fully solve the problem. The classification might be misled without considering the relationships among

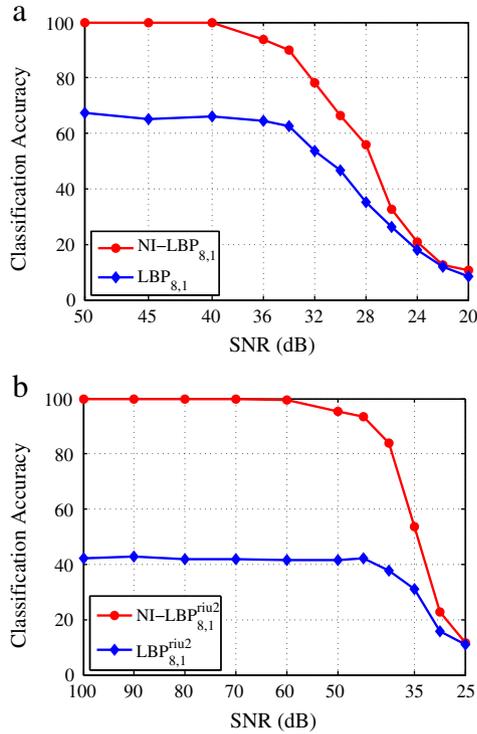


Fig. 5. Comparison of the robustness to additive Gaussian noise of different signal-to-noise ratios (SNR) for the proposed NI-LBP and the conventional LBP on the Outex textures: (a) $NI-LBP_{8,1}$ vs. $LBP_{8,1}$; (b) $NI-LBP_{8,1}^{riu2}$ vs. $LBP_{8,1}^{riu2}$. We have used all the original texture images present in the Outex_TC_00010 training set (20 samples of illuminant “inca” and angle 0 in each of the 24 texture classes, totaling 480 images). Training is done with all the 480 noise free images and testing is done with the same images, but added with additive Gaussian noise with different SNR. The nearest neighbor classifier is used for classification.

neighborhood intersection. Taking the patterns in Fig. 4(b) and (c) for example, which would be considered as the same pattern type according to LBP and VAR, they are actually two patterns with different textural properties. Moreover, MBP also fails to distinguish patterns (b) and (c). Clearly, our proposed NI-LBP approach can distinguish all the three different patterns, as shown in (a2) (b2) and (c2). Therefore, the proposed NI-LBP approach is more discriminative and effective.

In order to make further comparisons, we conducted texture classification on test suite Outex_TC_00000, which was used in [17]. The results are listed in Table 1. For test suite Outex_TC_00000, there are

24 texture classes, with each class having 20 monochrome texture images (128×128) with incandescent constant illumination and a spatial resolution 100 dpi. The images in each class are divided into two nonoverlap groups with 10 images as training and the other 10 as testing. Outex_TC_00000 provides 100 couples of test and train files for this category, we have performed experiments on all the 100 couples and we report classification accuracy as the average over the 100 couples. Note that in [17], the authors selected randomly one couple in their evaluation, see Table 1 in [17]. From Table 1, we can clearly observe that our NI-LBP performs the best, especially for two texture classes *canvas033* and *carpet009*.

Furthermore, LBP thresholding at the value of the central pixel $x_{0,0}$ tends to be sensitive to noise, particularly in near-uniform image regions, and smooths weak illumination gradients. While MBP thresholding against the median value is claimed to be robust to “salt and pepper” noise [17]. However, MBP is not robust to Gaussian noise. In contrast, the proposed NI-LBP descriptor has the following advantages:

1. Thresholding at μ is equivalent to making the local neighborhood vector zero-mean, therefore resistant to local lighting effects, and specifically invariant to gray scale changes.
2. Compared with LBP, weak edges are preserved by NI-LBP, as illustrated in Fig. 3. We can clearly observe that LBP does not match the visual patterns, producing output unrelated to the peak in (a) or the edge in (b). In contrast, the proposed NI-LBP outputs more consistent patterns, owing to the better thresholding of μ .
3. Better noise robustness, as shown in Fig. 5.

Recall that the local contrast measure proposed by Ojala et al. [4] is defined as follows:

$$VAR_{p,r} = \frac{1}{2} \sum_{n=0}^{p-1} (x_{r,n} - \mu)^2, \quad \text{where } \mu = \frac{1}{2} \sum_{n=0}^{p-1} x_{r,n}. \quad (7)$$

We can see that $NI-LBP_{p,r}$ and $VAR_{p,r}$ capture similar types of texture information, with slight differences:

1. $VAR_{p,r}$ achieves rotation invariance by summing up the whole variation in the circular neighborhood, whereas $NI-LBP_{p,r}$ is rotation sensitive, by default;
2. $NI-LBP_{p,r}$ is independent of gray scale, whereas $VAR_{p,r}$ is not;
3. Finally, $VAR_{p,r}$ is continuous-valued and needs to be quantized.

The latter quantization step has associated limitations of additional training to determine threshold values, and the difficulty in setting the number of bins. Too few bins will fail to provide enough discriminative information while too many bins would make the feature size too

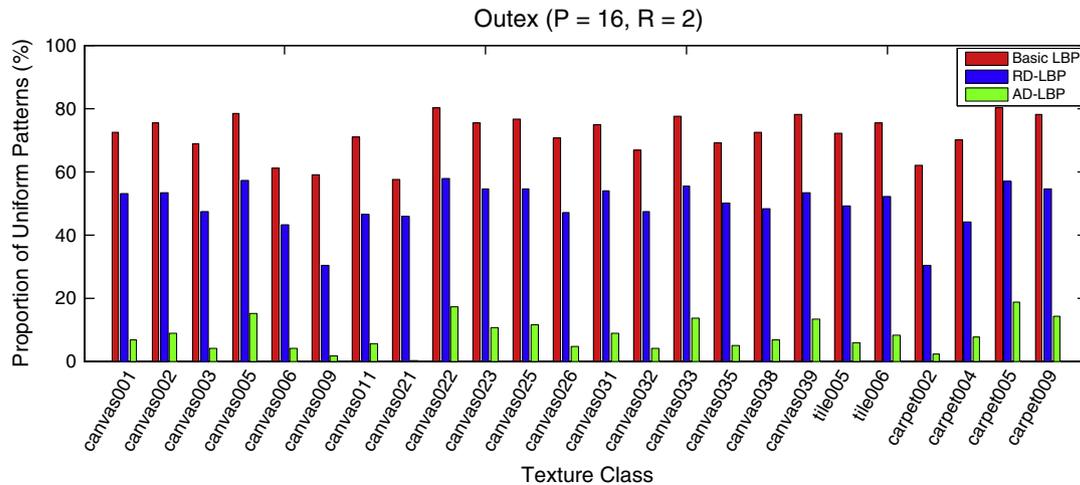


Fig. 6. Comparing the proportions (%) of “Uniform” patterns of all patterns for each texture in Outex for three methods: LBP, RD-LBP and AD-LBP, with $P=16, R=2$.

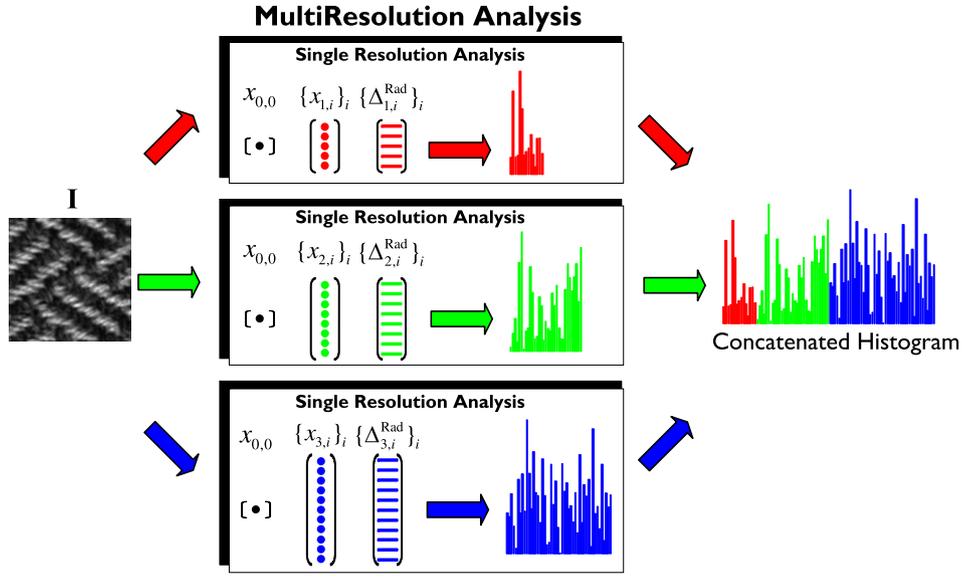


Fig. 7. Proposed multi-resolution scheme.

large. Although there are some rules to guide selection [4], it is hard to obtain an optimal number of bins in terms of accuracy and feature size. On the basis of the above discussion, we expect that the proposed $NI-LBP_{p,r}$ will be a better choice over $VAR_{p,r}$.

To make it consistent with the binary coding strategy, the 1D distribution of the central pixels' intensity is represented by two bins, i.e.,

$$CI-LBP = s(x_{0,0} - \mu_r), \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (8)$$

where μ_r is the mean of the whole image.

2.3. Difference-based descriptors

As a parallel development to the intensity descriptors just developed, we also propose pixel *differences* in radial and angular directions on a circular grid, different from the traditional pixel differences which are computed in horizontal and vertical directions. More specifically, we propose two different descriptors, Radial Difference Local Binary Pattern and Angular Difference Local Binary Pattern

(denoted as RD-LBP and AD-LBP respectively, as illustrated in Fig. 2). We define the RD-LBP descriptor as follows:

$$RD-LBP_{p,r,\delta} = \sum_{n=0}^{p-1} s(\Delta_{\delta,n}^{Rad}) 2^n, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (9)$$

where $\Delta_{\delta,n}^{Rad} = x_{r,n} - x_{r-\delta,n}$ is the radial difference computed with given integer radial displacement δ , $x_{r,n}$ and $x_{r-\delta,n}$ correspond to the gray values of pairs of pixels of δ equally spaced pixels of the same radial direction.

Similarly, the AD-LBP descriptor is defined as

$$AD-LBP_{p,r,\delta,\varepsilon} = \sum_{n=0}^{p-1} s(\Delta_{\delta,n}^{Ang}) 2^n, \quad s(x) = \begin{cases} 1, & x \geq \varepsilon \\ 0, & x < \varepsilon \end{cases} \quad (10)$$

where $\Delta_{\delta,n}^{Ang} = x_{r,n} - x_{r,\text{mod}(n+\delta,p)}$ is the angular difference computed with given angular displacement $\delta(2\pi/p)$, where δ is an integer such that $1 \leq \delta \leq p/2$, $x_{r,n}$ and $x_{r,\text{mod}(n+\delta,p)}$ correspond to the gray values of pairs of pixels of δ equally spaced pixels on a circular radius r , and function $\text{mod}(x,y)$ is the modulus x of y . ε is a threshold value,

Table 2

Summary of texture datasets used in our experiments.

Experiment # 1								
Texture dataset	Texture classes	Samples per class	Sample size	Test suite	Training or testing	Number of angles	Illuminant used	Samples in total
Brodatz	16	8	180 × 180	Contrib TC 00001 (problem 000–009)	Training	1	“inca”	16
					Testing	9	“inca”	1008
Outex	24	20	127 × 128	Outex TC 00010	Training	1	“inca”	480
					Testing	9	“inca”	3840
				Outex TC 00012 (problem 000)	Training	1	“inca”	480
					Testing	10	“tl84”	4320
				Outex TC 00012 (problem 001)	Training	1	“inca”	480
					Testing	10	“horizon”	4320
Experiment # 2								
Texture dataset	Dataset notation	Image rotation	Controlled illumination	Scale variation	Texture classes	Sample size	Samples per class	Samples in total
CURet	\mathcal{D}^c	✓	✓		61	200 × 200	92	5612
KTH-TIPS2b	\mathcal{D}^{KT2b}		✓	✓	11	200 × 200	432	4752

Table 3
Abbreviations for the notations of methods.

Name of Proposed method	Abbreviation
$LBP_{p,r}^{riu2}$	LBP
$VAR_{p,r}$	VAR
$LBP_{p,r}^{riu2}/VAR_{p,r}$	LBP/VAR
CI-LBP	CI
$NI-LBP_{p,r}^{riu2}$	NI
$RD-LBP_{p,r}^{riu2}$	RD
$RD-LBP_{p,r}^{riu2}/CI-LBP$	RD/CI
$NI-LBP_{p,r}^{riu2}/CI-LBP$	NI/CI
$NI-LBP_{p,r}^{riu2}/RD-LBP_{p,r}^{riu2}$	NI/RD
$NI-LBP_{p,r}^{riu2}/RD-LBP_{p,r}^{riu2}/CI-LBP$	NI/RD/CI

and is 1% of the pixel value range in our experiments. We experimentally set $\varepsilon = 0.01$. We can see that when $\delta = p/2$, our descriptor $AD-LBP_{p,r,p/2,\varepsilon}$ is equivalent to the CS-LBP descriptor proposed by Heikkilä [22] for local interesting region description.

As discussed in Section 2.1, limiting attention to the uniform binary patterns has attractive and elegant advantages over using all the binary patterns, specifically that the uniform patterns represent meaningful and fundamental characteristics of the texture, they appear to be the major parts of all binary patterns and are relatively reliable, and they lead to texture image representation of low dimensionality. We wish to see to what extent the proposed descriptors maintain these properties; Fig. 6 shows the proportions of the uniform patterns for three different descriptors (Basic LBP, RD-LBP and AD-LBP), extracted from texture images of test suite Outex_TC_00010. It can clearly be seen that the proportions of the uniform patterns of AD-LBP are too small to provide a reliable and meaningful description of texture. Consequently, we have decided against including AD-LBP descriptor in further experiments in this paper, and focus instead on RD-LBP.

Based on the above analysis, in order to produce acceptable dimensionality of histogram features, we merely use the uniform patterns motivated by the work of Ojala et al. [4]. The uniform patterns produce low dimensionality features so that they can be conveniently used together for pattern classification. There are two ways to combine the NI-LBP and RD-LBP codes: calculating the histograms separately and concatenating, or jointly, calculating a joint two dimensional histogram of the NI-LBP and RD-LBP codes, represented as NI-LBP/RD-LBP. Following the work of Varma and Zisserman [3] and Guo et al. [21], who showed the joint

approach to produce better results, we will prefer joint histogramming (shown in Fig. 2). Following [4], we use only joint distributions of operators that have the same (p, r) values, although nothing would prevent us from using joint distributions of operators computed from different neighborhoods.

2.4. Multiresolution analysis and classification

The proposed descriptors described above are extracted from a single resolution with a circularly symmetric neighbor set of p pixels placed on a circle of radius r (as in Fig. 2). Now clearly by altering (p, r) , we can realize operators for any quantization of the angular space and for any spatial resolution. Motivated by the idea of [4], we conduct the multiresolution analysis by combining the information provided by multiple descriptors of varying (p, r) , as illustrated in Fig. 7. The histogram feature vector of multiresolution analysis is obtained by concatenating the histograms from a single resolution analysis realized with different (p, r) .

To perform the actual texture classification, there are two crucial components: (i) texture feature extraction, and (ii) the classifier and the associated similarity measure used within the classifier. In this work the focus is on evaluating the discrimination properties of the proposed descriptors, so for classification we wish to make as few assumptions as possible and have chosen a non-parametric technique, since non-parametric classifiers can handle a large number of classes, avoid parameter overfitting, and require no learning/training. Of non-parametric classifiers, the k nearest neighbor (k NN) is one of the most popular and simplest methods, which we adopt with $k = 1$. The samples are then classified according to their normalized histogram feature vectors \underline{h}_i and \underline{h}_j , using χ^2 distance metric

$$\chi^2(\underline{h}_i, \underline{h}_j) = \frac{1}{2} \sum_k \frac{[\underline{h}_i(k) - \underline{h}_j(k)]^2}{\underline{h}_i(k) + \underline{h}_j(k)} \quad (11)$$

the same distance metric used in [2, 3, 27, 30].

3. Experimental evaluation

In this section, we demonstrate the performance of the proposed method with comprehensive experiments on six texture datasets, summarized in Table 2, which are derived from four popular publicly

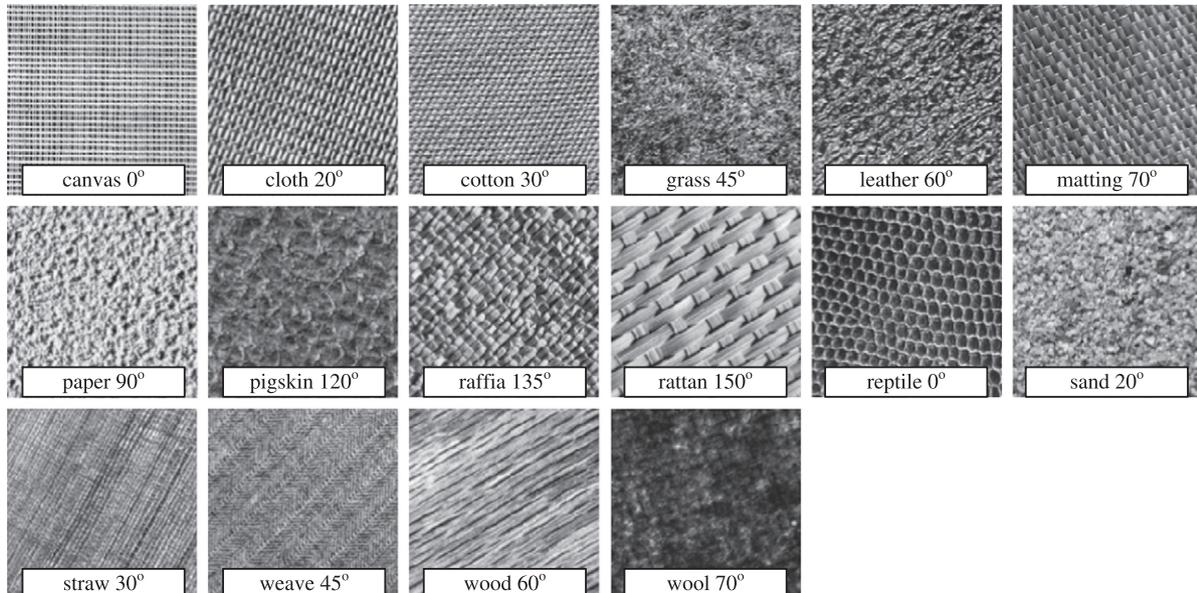


Fig. 8. 128×128 samples of the textures from Brodatz used in Experiment #1.

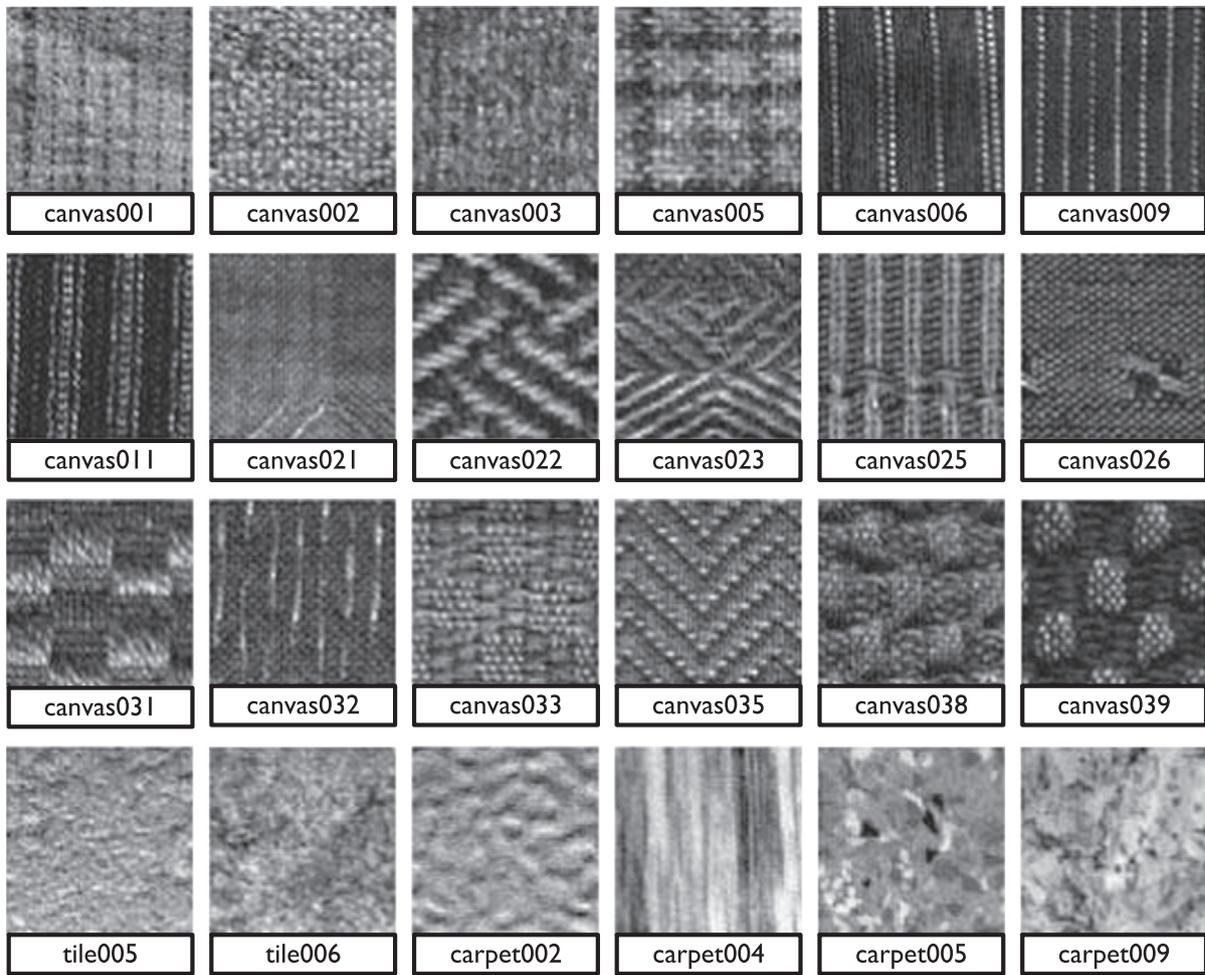


Fig. 9. 128 × 128 samples of the 24 textures from Outex used in Experiment #1.

Table 4

Classification accuracies (%) on Contrib_TC_00001, where training is done at just one rotation angle and the average accuracy over 10 angles. The results for LBP, VAR, and LBP/VAR are quoted directly from the original paper by Ojala et al. [4].

2 * Method	2 * (p, r)	2 * Bins	Rotation angle for training										2 * Average
			0°	20°	30°	45°	60°	70°	90°	120°	135°	150°	
LBP	(16, 2)	18	96.2	99.0	98.6	98.9	98.5	99.1	97.6	98.6	98.7	97.5	98.3
VAR	(16, 2)	128	89.9	84.5	986.2	90.5	87.3	85.6	91.0	89.8	90.8	88.5	88.4
LBP/VAR [4]	(8, 1) + (16, 2) + (24, 3)	864	100	99.7	99.5	99.8	99.6	99.7	99.8	99.6	99.8	99.9	99.7
2 * NI	(8, 1)	10	65.4	85.5	81.3	76.6	77.0	78.4	68.8	81.4	75.8	76.5	76.7
	(16, 2)	18	87.6	95.2	92.3	93.6	89.4	96.0	88.9	91.3	93.4	90.1	91.8
	(24, 3)	26	96.2	93.4	97.6	96.6	98.3	96.7	97.1	96.7	92.6	98.2	96.4
2 * RD	(8, 1)	10	68.8	86.4	84.4	76.0	84.9	84.4	70.2	84.1	76.1	84.7	80.0
	(16, 2)	18	89.2	92.9	96.7	97.8	96.1	92.6	88.4	94.7	96.7	97.3	94.3
	(24, 3)	26	87.6	90.6	98.2	90.8	96.5	93.8	89.5	98.6	89.5	94.2	92.9
2 * RD/CI	(8, 1)	20	87.1	84.7	94.3	88.6	95.9	95.1	85.8	94.8	90.3	95.0	92.2
	(16, 2)	36	92.7	94.6	96.8	97.3	98.4	95.6	91.8	99.4	96.7	98.6	96.2
	(24, 3)	52	96.9	95.8	95.6	92.8	96.5	94.3	96.9	99.1	95.3	95.9	95.9
2 * NI/CI	(8, 1)	20	74.8	90.4	86.4	80.3	82.5	85.2	74.4	86.2	80.6	82.2	82.2
	(16, 2)	36	95.6	99.2	98.8	98.0	98.2	99.4	93.8	98.3	96.9	97.4	97.6
	(24, 3)	52	99.1	98.7	99.4	99.4	100	100	99.7	97.5	97.3	99.1	99.1
2 * NI/RD/CI	(8, 1)	100	70.2	88.9	87.0	80.0	85.2	85.5	71.9	87.1	81.6	84.9	82.2
	(16, 2)	324	100	100	100	100	100	100	100	100	100	100	100
	(24, 3)	676	98.2	100	100	100	100	100	99.6	99.9	99.9	100	99.8
2 * 100.0%	(8, 1)	200	78.1	94.5	92.2	91.1	93.0	92.0	76.2	92.4	91.8	92.6	89.4
	(16, 2)	648	100	100	100	100	100	100	100	100	100	100	100
	(24, 3)	1352	98.8	100	100	100	100	100	99.8	100	99.8	100	99.8

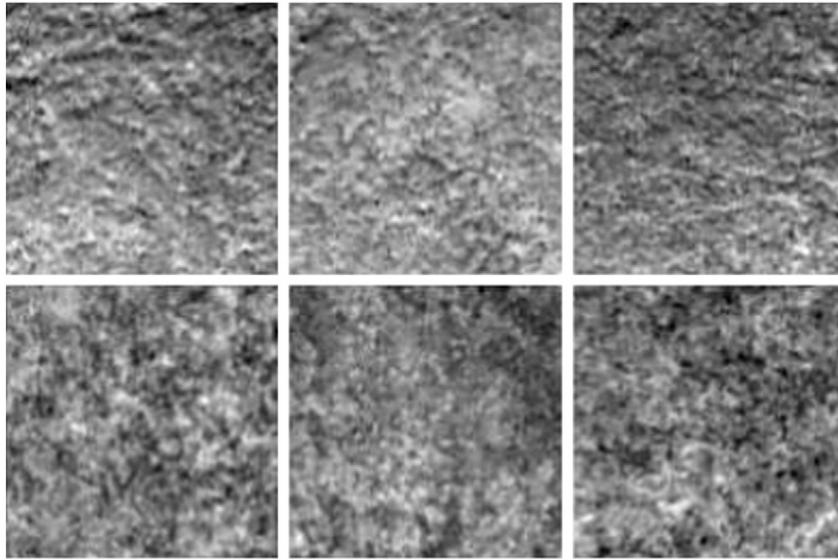


Fig. 10. Some example texture samples from tile005 (top row) and tile006 (bottom row). We can see that they look fairly similar.

abbreviations are summarized in Table 3. In all experiments, each texture sample is normalized to be zero mean and unit standard deviation. Results for the CURET database are reported over 100 random partitions of training and testing sets. 1NN is used for classification.

3.2. Experiment #1

3.2.1. Image data and experimental setup

Contrib_TC_00001: This test suite consists of 16 texture classes from the Brodatz database [31] (a few shown in Fig. 8). This test suite was designed for rotation invariant texture classification note [2] http://www.ee.oulu.fi/mvg/page/image_data. There are eight samples of size 180×180 in each class, out of which the first sample is utilized for training and the other seven as testing. Given ten rotation angles, the classifier is trained with samples artificially rotated to just one angle and tested against samples rotated to the other nine angles. In each experiment, the classifier was trained with 16 images and tested with 1008 ($16 \times 7 \times 9$) samples, 63 in each of the 16 texture classes.

Following [4], each training sample is split into 121 disjoint 16×16 sub-samples, whose histograms are then merged into one model histogram. We point out that the seven testing images in each texture class are physically different from the one designated training image.

Outex_TC_00010: 24 Outex texture classes (shown in Fig. 9) with each class having 20 samples. It was created by Ojala et al. [4], again for rotation invariant texture classification. All textures in this test suite have the same illuminant “inca”. The training and testing scheme is the same as that for Contrib_TC_00001 but with nine different rotation angles. All of the 480 (24×20) samples rotated by one angle are adopted as the training data, and testing data consists of all 480 samples rotated by the other 8 angles. Hence, there are 480 models for training, and 3840 (480×8) for validation.

Outex_TC_00012: Created by Ojala et al. [4] for rotation and illumination invariant texture classification. The texture classes are the same as Outex_TC_00010. The classifier was trained with the same training samples as Outex_TC_00010, but tested with all samples captured at all 9 rotation angles under different illuminants “t184” or “horizon”. Due

Table 7

Classification accuracies (%) of descriptor NI/RD/CI for Outex_TC_00010 and Outex_TC_00012: training is done at just one rotation angle, and the average accuracy over 9 angles.

Test suite	(p, r)	Rotation Angle for Train (“inca”)									
		0°	5°	10°	15°	30°	45°	60°	75°	90°	Average
7*[c] Outex_ TC_00012 (“t184”)	(8, 1)	90.9	91.6	92.1	93.0	91.3	90.8	88.9	89.0	84.3	90.2
	(16, 2)	98.0	98.3	99.1	98.6	98.4	98.6	98.6	97.7	96.8	98.3
	(24, 3)	97.3	98.3	98.5	98.7	97.2	96.4	93.4	94.2	94.1	96.5
	(8, 1) + (16, 2)	97.4	98.0	98.4	98.5	98.3	98.3	97.8	97.1	95.6	97.7
	(8, 1) + (24, 3)	97.7	93.3	98.7	98.7	98.5	97.9	96.4	96.6	96.4	97.7
	(16, 2) + (24, 3)	98.3	99.0	99.3	99.2	98.9	98.9	98.3	98.1	98.1	98.7
	(8, 1) + (16, 2) + (24, 3)	98.5	98.9	99.1	99.1	99.0	98.9	98.4	98.2	98.1	98.7
7*[c] Outex_ TC_00012 (“horizon”)	(8, 1)	92.7	92.8	93.3	93.6	92.7	91.6	90.3	91.1	86.6	91.6
	(16, 2)	98.0	98.0	98.3	98.4	97.7	97.9	98.2	98.3	98.1	98.1
	(24, 3)	96.2	97.0	97.0	97.3	95.5	95.1	92.7	93.7	94.1	95.4
	(8, 1) + (16, 2)	98.2	97.8	98.3	97.9	97.1	97.8	98.2	97.8	97.0	97.8
	(8, 1) + (24, 3)	97.8	97.5	97.7	97.7	96.2	96.1	95.1	95.2	95.1	96.3
	(16, 2) + (24, 3)	97.8	98.3	98.2	98.3	97.3	97.5	96.9	97.0	97.7	97.7
	(8, 1) + (16, 2) + (24, 3)	97.8	98.4	98.4	98.2	97.4	97.7	97.5	97.1	97.6	97.8
7*[c] Outex_ TC_00010 (“inca”)	(8, 1)	96.5	96.3	97.4	97.6	96.2	95.3	92.7	94.9	91.8	95.4
	(16, 2)	99.3	99.4	99.5	99.7	99.6	99.6	99.5	99.0	99.0	99.4
	(24, 3)	99.2	99.5	99.4	99.5	99.5	99.5	99.2	99.3	99.1	99.4
	(8, 1) + (16, 2)	99.4	99.4	99.6	99.6	99.5	99.4	99.4	99.0	98.6	99.3
	(8, 1) + (24, 3)	99.3	99.5	99.5	99.5	99.6	99.6	99.7	99.4	99.2	99.5
	(16, 2) + (24, 3)	99.6	99.7	99.8	99.7	99.7	99.9	99.8	99.7	99.5	99.7
	(8, 1) + (16, 2) + (24, 3)	99.7	99.7	99.7	99.6	99.6	99.8	99.9	99.7	99.4	99.7

The bold numbers indicate the highest classification score achieved on each dataset.

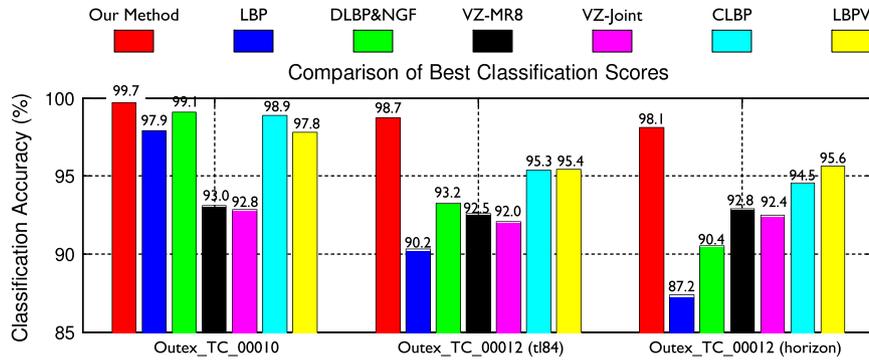


Fig. 11. Comparing the best classification scores of our approach with various state-of-the-art methods on all the three test suites. All the results are as originally reported, except for those of VZ-MR8 and VZ-Joint, which are obtained by us using the exact same experimental setup as Varma and Zisserman did [2, 3]. For VZ-MR8 and VZ-Joint, 40 textons per class is used for building the universal texton dictionary.

to the varying illuminants, some texture samples have a large tactile dimension which induces significant local gray-scale distortions, therefore Outex_TC_00012 is more challenging than Outex_TC_00010.

3.2.2. Experimental results on Contrib_TC_00001

Ojala et al. [4] reported a near-perfect classification accuracy of 99.7% for the joint descriptor *LBP VAR* when using two spatial resolutions (8,1) + (24,3) or three spatial resolutions (8,1) + (16,2) + 24,3. Table 4 presents the results for our proposed descriptors, comparing with the state-of-the-art methods [4].

The individual descriptor *NI-LBP* and *RD-LBP* perform similarly, with *NI-LBP* doing slightly better. $NI-LBP_{16,2}^{riu_2}$ and $NI-LBP_{24,3}^{riu_2}$

significantly outperformed their simpler counterpart $NI-LBP_{8,1}^{riu_2}$. This is also the case with *RD-LBP*. Interestingly, the performance of *NI-LBP* increases with the neighborhood size, while for *RD-LBP*, the best performance is achieved by $RD-LBP_{16,2}^{riu_2}$. On average, between the individual descriptors, *LBP* performs the best and *VAR* the worst.

The center pixel also provides useful discriminative information, since it is apparent in Table 4 that combining the center pixel *CI-LBP* with *NI-LBP* or *RD-LBP* can generally improve classification performance. Neglecting the center pixel clearly results the loss in information, similar to how [3] and [21] demonstrated the benefits of explicitly including the information of the center pixel in the classifier.

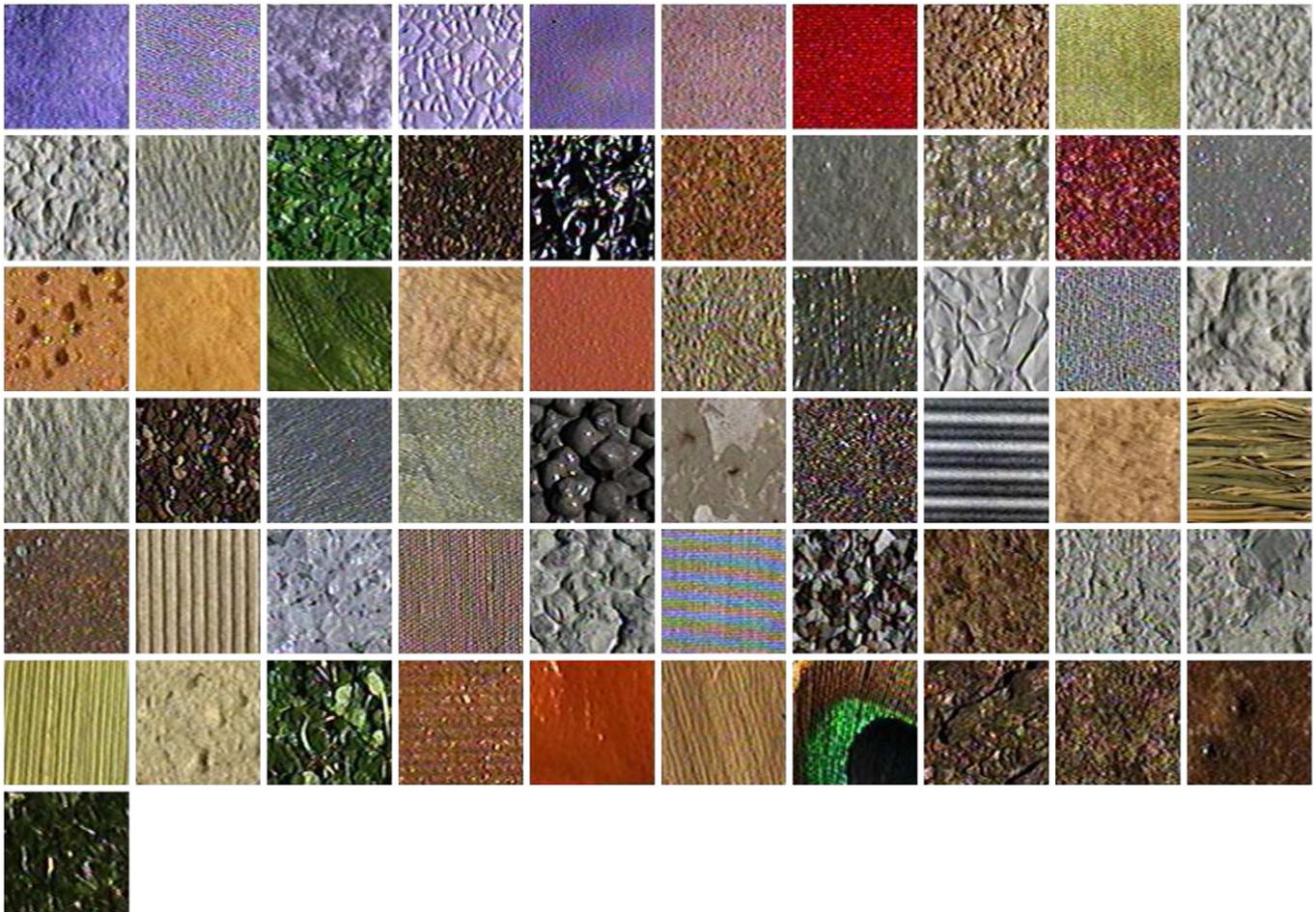


Fig. 12. One sample of each of the 61 texture classes from CURET.



Fig. 13. The variations within each category of the new KHTIPS2b database. Each row shows one example image from each of four samples of a category.

The fusion of *NI-LBP* and *RD-LBP* produced perfect classification results at (16,2). The much improved classification accuracy by combining *NI-LBP* and *RD-LBP*, which were on their own no better than *LBP*, implies that they capture truly complementary texture information.

It is evident, based on the results, that the performance of the proposed *NI-LBP/RD-LBP* and *NI-LBP/RD-LBP/CI-LBP* descriptors are superior to that of *LBP/VAR*. Note that here we only consider a single resolution for our descriptors. It may be argued that this test suite is too easy for texture classification; more challenges of test suites follow.

3.2.3. Experimental results on Outex_TC_00010 and Outex_TC_00012

Table 5 presents results for our proposed descriptors and those proposed in [4] on test suites Outex_TC_00010 and Outex_TC_00012. The conclusions from these results are similar to those from Table 4, with the following additional observations.

First, our proposed descriptor $NI - LBP_{16,2}^{riu2}/RD - LBP_{16,2}^{riu2}/CI - LBP$ produces consistently the best classification scores across all three test suites, a considerable improvement over the best reported results by Ojala et al. [4], especially for Outex_TC_00012 “t184” and “horizon”.

Second, among individual descriptors, although *NI-LBP* and *RD-LBP* did not outperform *LBP* or *VAR*, their combination significantly outperformed *LBP/VAR*. We maintain that the *NI-LBP/RD-LBP* strength stems from their complementarity, in that *NI-LBP* measures the variation of the neighboring pixels on the same circumference, while *RD-LBP* captures the edge information between circumferences, analogous to the combination of RIFT and SIFT used by Zhang et al. [8].

Finally, we can see that $NI - LBP_{16,2}^{riu2}/RD - LBP_{16,2}^{riu2}$ and $NI - LBP_{16,2}^{riu2}/RD - LBP_{16,2}^{riu2}/CI - LBP$ produce very robust classification performance in all three cases. This is in contrast to the descriptor *LBP/VAR*, the performance of which decreases considerably in gray scale and rotation invariant texture classification. The excellent classification results demonstrate that *NI-LBP/RD-LBP/CI-LBP* is more stable for texture classification irrespective of the different imaging geometries of the illuminants affecting the appearance of local distortions caused by the tactile dimension of textures.

Table 6 shows the number of misclassified samples for each texture and rotation angle for the best descriptor $NI - LBP_{16,2}^{riu2}/LBP - R_{16,2}^{riu2}/CI - LBP$ for all three cases, allowing a detailed analysis of discrimination of individual textures and the effect of rotation. Overall, $NI - LBP_{16,2}^{riu2}/RD - LBP_{16,2}^{riu2}/CI - LBP$ classified 18 out of the 24 classes completely correctly, having most difficulties with *tile006*, followed by *canvas033*.

Incidentally, for all the three test suites nearly all of the misclassified samples in *tile006* and *tile005* were assigned to each other. Fig. 10 shows some example textures from class *tile005* and *tile006*, where we can observe the high degree of perceptual similarity.

Motivated by its excellent classification performance and in order to fully examine the classification performance of descriptor *NI-LBP/RD-LBP/CI-LBP*, Table 7 shows the results of extensive experiments which we conducted on the three Outex test suites by varying the training angle. We can see that the performance is very robust, especially true with $NI - LBP_{16,2}^{riu2}/RD - LBP_{16,2}^{riu2}/CI - LBP$. We can also observe the better results obtained by multiresolution analysis over single resolution. We acknowledge that the multiresolution analysis

Table 8

Comparing classification accuracy (%) on CURET: N_{tr} is the number of training samples per class used. All results are obtained by us except for VZ-Joint, which are quoted from the recent comparative study of Zhang et al. [8]. For VZ-MR8, we learn 10 textons per class.

(p, r)	(8, 1)	(16, 2)					(24, 3)								
N_{tr}	46	23	12	6	2	46	23	12	6	2	46	23	12	6	2
<i>LBP/VAR</i>	93.76	88.71	81.80	71.08	50.43	4.00	89.76	81.53	71.09	52.77	91.90	85.34	77.12	66.04	48.64
<i>NI/RD/CI</i>	95.15	92.00	86.19	77.97	57.96	9563	92.7	87.12	79.57	60.89	92.59	87.85	80.92	70.33	52.21
(p, r)	(8, 1) + (16, 2)	(8, 1) + (24, 3)					(8, 1) + (16, 2) + (24, 3)								
N_{tr}	46	23	12	6	2	46	23	12	6	2	46	23	12	6	2
<i>NI/RD</i>	94.78	91.17	85.73	76.67	57.28	95.79	91.17	86.26	77.25	55.68	95.75	91.88	85.41	75.71	57.80
<i>NI/RD/CI</i>	96.88	93.55	89.29	80.18	61.28	96.66	93.57	88.41	79.90	60.52	96.78	93.45	88.94	79.69	62.14
Method	Neighborhood size	46	23	12	6	2									
VZ-MR8	19 × 19	96.37	92.34	86.96	77.17	54.88									
VZ-Joint	7 × 7	96.19	92.00	86.56	76.87	54.69									

Table 9

Comparison of highest classification performance on CURET with state-of-the-art results using NNC classifier. Number of training and testing samples per class is equal, i.e. 46. Our score 97.29% is obtained with *NI/RD/CI* at multiresolutions (8, 1) + (16, 2) + (24, 5). All the results from other methods are quoted directly from the original papers except for those of *LBP/VAR*, which are obtained by us.

Method	Ours	LBP/ VAR	VZ-MR8 [2]	VZ-Joint [3]	VZ-MRF [3]	CLBP [21]
Neighborhood size	11×11	5×5	49×49	19×19	11×11	7×7
Bins	2200	416	2440	610	219,600	2200
Accuracy (%)	97.29	94.00	97.43	97.17	98.03	97.39

will increase the dimensionality of the histogram feature, however the largest dimensionality of 2200 for three resolutions is not a big problem.

To conclude Experiment #1, Fig. 11 compares the best scores achieved by our proposed method and those reported by six other state-of-the-art methods. It is quite clear that our approach consistently outperforms all state-of-the-art methods in gray scale and rotation invariant texture classification.

It is important to emphasize that although our proposed descriptors are motivated by LBP, in practice we are extracting very different local texture information, whereas DLBP, LBPV and CLBP are all LBP-based approaches. LBP, CLBP, and our proposed approach share the advantage of being training-free and computationally simple, since they are based upon a pre-defined dictionary rather than one derived with reference to the dataset to be analyzed. In contrast, VZ-MR8 and VZ-Joint require a time-consuming universal texton dictionary learning stage by clustering local feature vectors extracted from training samples. From Fig. 11 we see that our approach has about a 5%–7% improvement over VZ-MR8 and VZ-Joint, most likely in part due to the limited training samples for learning the universal texton dictionary, leading to a drop in accuracy for VZ-MR8 and VZ-Joint.

3.3. Experiment #2

Motivated by the excellent performance for the proposed approach demonstrated in the previous section, here we test the performance of the proposed approach for material classification and categorization, using the CURET and KTH TIPS2b databases.

3.3.1. Image data and experimental setup

CURET [2, 3, 35]: The original CURET database [35] consists of 61 texture classes, shown in Fig. 12, with each class containing 205 images of a physical texture sample photographed under a (calibrated) range of viewing and lighting angles, but without significant variation

Table 10

Classification results (%) of the proposed descriptors and the LBP on KTH TIPS2b. $CLBP_S/M/C$ represents $CLBP_S^{riu2}/M_{p,r}^{riu2}/C$.

(p, r)	(8, 1)			(16, 2)			(24, 3)		
	1	2	3	1	2	3	1	2	3
LBP	48.1	54.2	56.8	50.5	55.8	59.1	49.9	54.6	57.8
NI	46.1	48.7	52.3	37.6	41.2	44.4	39.0	42.6	45.4
RD	48.1	54.2	56.9	44.0	49.4	52.1	38.5	42.7	45.2
RD/CI	52.6	57.8	61.2	49.9	56.2	59.8	48.0	53.5	56.3
NI/CI	47.1	53.1	56.4	46.5	50.5	53.0	44.8	48.8	51.6
NI/RD	53.5	60.0	63.1	54.0	59.0	61.6	49.5	55.2	59.1
NI/RD/CI	56.6	61.9	64.8	57.7	62.5	65.1	52.4	57.5	61.7
(p, r)	(8, 1) + (16, 2)			(16, 2) + (24, 3)			(8, 1) + (16, 2) + (24, 3)		
N _{train}	1	2	3	1	2	3	1	2	3
NI/RD/CI	58.1	62.9	66.0	55.9	61.0	64.2	56.7	61.7	65.0

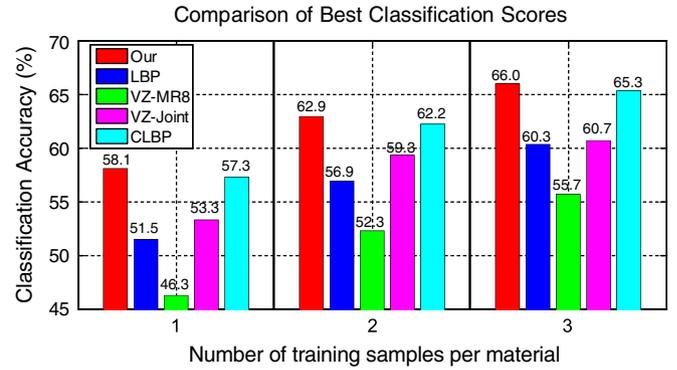


Fig. 14. Comparing the proposed approach with various state-of-the-art methods on KTH TIPS2b.

in scale or in-plane rotation. CURET is a challenging test of texture descriptors because of the large intra-class variation including the effects of specularities, interreflections, shadowing, and other surface normal variations due to lighting geometry. Consistent with other CURET studies [2, 3], we consider only the 92 images per class which afford the extraction of a 200×200 pixel foreground region of texture, the same subset of images as in [2, 3].

KTH TIPS2b [36, 27]: It is generally agreed [2, 3, 27] that the major drawback of the CURET database is that materials are imaged at a constant scale. The acquisition procedure for KTH TIPS2b has been described in more detail in [36], with 3 viewing angles, 4 illuminants, and 9 different scales, producing 432 images per class. Fig. 13 illustrates an example of the 11 materials. Notice in particular the striking differences between samples of the same class. There is almost no intra-class variation due to in-plane rotation for this database.

For the experiments on KTH TIPS2b, we follow the training and testing scheme used in [27]. We perform experiments training on one, two, or three samples; testing is always conducted only on unseen samples.

3.3.2. Experimental results

Table 8 presents the results on CURET. Our method consistently outperformed *LBP/VAR*, and the multiresolution analysis of the proposed approach helps to improve classification performance, producing slightly higher classification scores than VZ-MR8 and VZ-Joint. Moreover, it is clear that the degree of improvement of our descriptor *NI-LBP/RD-LBP/CI-LBP* over that of *LBP/VAR* is increased given fewer training samples, in accordance with the findings based on the three Outex test suites.

In order to make the comparison fair, Table 9 compares the best classification scores achieved by various state-of-the-art methods on CURET. We can see that our proposed approach is outperformed by VZ-MR8, VZ-Joint, and VZ-MRF with large neighborhood size and more textons. This is because VZ-MR8 and VZ-Joint are statistical approaches, and the very large number of training samples in this dataset allows those methods to find representative textons; with reduced training data the performance of the VZ methods would decrease. Nevertheless, our method, despite a small spatial support, can compete with VZ-MR8 and VZ-Joint having a much larger spatial support. The lesser performance, in general, of the LBP methods is that there are scale and affine variations in the CURET database, while LBP-based approaches are proposed for rotation and gray level invariance and have limited capability to address scale and affine invariance.

We have conducted experiments with larger neighborhood sizes on CURET, and our descriptor *NI/RD/CI* at multiresolutions (8,1) + (16,2) + (24,5) gives classification scores of 97.29%, 94.48%, 88.96%, 80.70% for 46, 23, 12, and 6 training samples per texture class, respectively, which are slightly better than the results achieved by the CLBP approach (97.39%, 94.19%, 88.72%, and 79.88%, respectively).

As shown in Table 10 and Fig. 14, we also compare our method with state-of-the-art methods on the material categorization task of the KTH-TIPS2b textures, with all results from other methods quoted directly from [27]. For this database, our proposed NI-LBP/RD-LBP/CI-LBP descriptor outperforms all compared state-of-the-art methods by a significant margin. We should bear in mind that the classification results of all of the methods are obtained with a 1NN classifier, since we mainly focus our attention on the effectiveness of the descriptors rather than on the capabilities of the classifier. Using a more advanced classifier (SVM or $k > 1$) might improve performance significantly.

4. Conclusions and future work

This paper has proposed a novel local texture descriptor, generalizing the well-known LBP approach. Four LBP-like descriptors, two local intensity-based CI-LBP and NI-LBP, and two local difference-based descriptors RD-LBP and AD-LBP, were presented to extract complementary texture information of local spatial patterns. We showed that combining complementary descriptors played an important role in texture discrimination. In addition, we found that information contained in radial differences is more discriminative than those contained in angular difference.

The advantages of the proposed approach include its computational simplicity, no training (in the feature extraction stage), and a data-independent universal texton dictionary. Extensive experimental results show that the joint distribution of CI-LBP, NI-LBP and RD-LBP significantly outperform the conventional LBP approach and its various invariants on the Outex test suites. Furthermore, results on the material database KTH-TIPS2b demonstrate the best performance of the proposed approach in comparison with several state-of-the-art methods with a nearest neighbor classifier.

In the future, we plan to explore how to reduce the feature dimension of the multiresolution CI-LBP/NI-LBP/RD-LBP. We also believe that an in-depth investigation of the AD-LBP descriptor would be valuable for local region description, looking at the parallels between AD-LBP and the CS-LBP of [22] developed for image matching.

References

- [1] M. Tuceryan, A.K. Jain, Texture analysis, in: C.H. Chen, L.F. Pau, P.S.P. Wang (Eds.), Handbook Pattern Recognition and Computer Vision, World Scientific, Singapore, 1993, pp. 235–276.
- [2] M. Varma, A. Zisserman, A statistical approach to texture classification from single images, *Int. J. Comput. Vis.* 62 (1–2) (2005) 61–81.
- [3] M. Varma, A. Zisserman, A statistical approach to material classification using image patches, *IEEE Trans. Pattern Anal. Mach. Intell.* 31 (11) (2009) 2032–2047.
- [4] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (7) (2002) 971–987.
- [5] T. Leung, J. Malik, Representing and recognizing the visual appearance of materials using three-dimensional textons, *Int. J. Comput. Vis.* 43 (1) (2001) 29–44.
- [6] L. Zhang, L. Zhang, Z. Guo, D. Zhang, Monogenic-LBP: a new approach for rotation invariant texture classification, *IEEE International Conference on Image Processing (ICIP)*, 2010, pp. 2677–2680.
- [7] J. Xie, L. Zhang, J. You, D. Zhang, Texture classification via patch-based sparse texton learning, *IEEE International Conference on Image Processing (ICIP)*, 2010, pp. 2737–2740.
- [8] J. Zhang, M. Marszałek, S. Lazebnik, C. Schmid, Local features and kernels for classification of texture and object categories: a comprehensive study, *Int. J. Comput. Vis.* 73 (2) (2007) 213–238.
- [9] M. Crosier, L.D. Griffin, Using basic image features for texture classification, *Int. J. Comput. Vis.* 88 (3) (2010) 447–460.
- [10] S. Lazebnik, C. Schmid, J. Ponce, A sparse texture representation using local affine regions, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (8) (2005) 1265–1278.
- [11] B.S. Manjunath, W.Y. Ma, Texture features for browsing and retrieval of image data, *IEEE Trans. Pattern Anal. Mach. Intell.* 18 (8) (1996) 837–842.
- [12] T. Randen, J. Husøy, Filtering for texture classification: a comparative study, *IEEE Trans. Pattern Anal. Mach. Intell.* 21 (4) (1999) 291–310.
- [13] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures with classification based on feature distributions, *Pattern Recognit.* 29 (1) (1996) 51–59.
- [14] Y. Rodriguez, S. Marcel, Face authentication using adapted local binary pattern histograms, *European Conference on Computer Vision (ECCV)*, Graz, Austria, 2006, pp. 321–332.
- [15] T. Ahonen, A. Hadid, M. Pietikäinen, Face description with local binary patterns: application to face recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (12) (2006) 2037–2041.
- [16] G. Zhao, M. Pietikäinen, Dynamic texture recognition using local binary patterns with an application to facial expressions, *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (6) (2007) 915–928.
- [17] A. Hafiane1, G. Seetharaman, B. Zavidovique, Median binary pattern for textures classification, *Proceedings of ICIAR*, 2007, pp. 387–398.
- [18] M. Heikkilä, M. Pietikäinen, J. Heikkilä, A texture-based method for detecting moving objects, *British Machine Vision Conference (BMVC)*, London, vol. 1, 2004, pp. 187–196.
- [19] M. Pietikäinen, T. Nurmela, T. Mäenpää, M. Turtinen, View-based recognition of real-world textures, *Pattern Recognit.* 37 (2) (2004) 313–323.
- [20] S. Liao, M.W.K. Law, A.C.S. Chung, Dominant local binary patterns for texture classification, *IEEE Trans. Image Process.* 18 (5) (2009) 1107–1118.
- [21] Z. Guo, L. Zhang, D. Zhang, A completed modeling of local binary pattern operator for texture classification, *IEEE Trans. Image Process.* 9 (16) (2010) 1657–1663.
- [22] M. Heikkilä, M. Pietikäinen, C. Schmid, Description of interest regions with local binary patterns, *Pattern Recognit.* 42 (3) (2009) 425–436.
- [23] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikäinen, X. Chen, W. Gao, WLD: A robust local image descriptor, *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (9) (2010) 1705–1720.
- [24] G.K. Cross, A.K. Jain, Markov random field texture models, *IEEE Trans. Pattern Anal. Mach. Intell.* 5 (1) (1983) 25–39.
- [25] T. Ojala, K. Valkealahti, E. Oja, M. Pietikäinen, Texture discrimination with multi-dimensional distributions of signed gray-level differences, *Pattern Recognit.* 34 (3) (2001) 727–739.
- [26] L. Liu, P. Fieguth, G. Kuang, Generalized local binary patterns for texture classification, *British Machine Vision Conference (BMVC2011)*, 2011.
- [27] B. Caputo, E. Hayman, P. Mallikarjuna, Class-specific material categorization, *International Conference on Computer Vision (ICCV)*, Beijing, 2005, pp. 1597–1604.
- [28] T. Ahonen, M. Pietikäinen, Soft histograms for local binary patterns, *Finnish Signal Processing Symposium*, Oulu, Finland, 2007.
- [29] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, *IEEE Trans. Image Process.* 19 (6) (2010) 1635–1650.
- [30] B. Caputo, E. Hayman, M. Fritz, J.-O. Eklundh, Classifying materials in the real world, *Image Vis. Comput.* 28 (1) (2010) 150–163.
- [31] P. Brodatz, *Texture: A Photographic Album for Artists and Designers*, Dover, New York, 1966.
- [32] T. Ojala, T. Mäenpää, M. Pietikäinen, J. Viertola, J. Kyllönen, S. Huovinen, Outex—new frame work for empirical evaluation of texture analysis algorithm, *International Conference on Pattern Recognition*, 2002, pp. 701–706.
- [33] Z. Guo, L. Zhang, D. Zhang, Rotation invariant texture classification using LBP variance (LBPV) with global matching, *Pattern Recognit.* 43 (3) (2010) 706–719.
- [34] L. Nanni, A. Lumini, S. Brahmam, Local binary patterns variants as texture descriptors for medical image analysis, *Artif. Intell. Med.* 49 (2) (2010) 117–125.
- [35] K.J. Dana, B. van Ginneken, S.K. Nayar, J.J. Koenderink, Reflectance and texture of real-world surfaces, *ACM Trans. Graph.* 18 (1) (1999) 1–34.
- [36] P. Mallikarjuna, M. Fritz, A.T. Targhi, E. Hayman, B. Caputo, J.-O. Eklundh, The KTH-TIPS and KTH-TIPS2 Databases, <http://www.nada.kth.se/cvap/databases/kth-tips/> 2006.