Pattern Recognition 45 (2012) 2405–2418

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Sorted random projections for robust rotation-invariant texture classification

Li Liu [a,*], Paul Fieguth [b], David Clausi [b], Gangyao Kuang [a]

[a] School of Electronic Science and Engineering, National University of Defense Technology, 47 Yanwachi, Changsha 410073, Hunan, China
[b] Department of System Design Engineering, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1

ARTICLE INFO

Article history:
Received 30 November 2010
Received in revised form
1 May 2011
Accepted 26 October 2011
Available online 23 December 2011

Keywords:
Texture classification
Random projection
Compressed sensing
Feature extraction
Image patches
Rotation invariance
Bag of words

ABSTRACT

This paper presents a simple, novel, yet very powerful approach for robust rotation-invariant texture classification based on random projection. The proposed sorted random projection maintains the strengths of random projection, in being computationally efficient and low-dimensional, with the addition of a straightforward sorting step to introduce rotation invariance. At the feature extraction stage, a small set of random measurements is extracted from sorted pixels or sorted pixel differences in local image patches. The rotation invariant random features are embedded into a bag-of-words model to perform texture classification, allowing us to achieve global rotation invariance. The proposed unconventional and novel random features are very robust, yet by leveraging the sparse nature of texture images, our approach outperforms traditional feature extraction methods which involve careful design and complex steps. We report extensive experiments comparing the proposed method to six state-of-the-art methods, RP, Patch, LBP, WMFS and the methods of Lazebnik et al. and Zhang et al., in texture classification on five databases: CUReT, Brodatz, UIUC, UMD and KTH-TIPS. Our approach leads to significant improvements in classification accuracy, producing consistently good results on each database, including what we believe to be the best reported results for Brodatz, UMD and KTH-TIPS.

1. Introduction

Texture is an important characteristic of the appearance of objects in natural scenes and is a powerful visual cue, used by both humans and machines in describing and recognizing real world object surfaces. Texture analysis is an active research area spanning image processing, pattern recognition, and computer vision, with applications to medical image analysis, remote sensing, object recognition, industrial surface inspection, document segmentation and content-based image retrieval. Texture classification has received significant attention with many proposed approaches, as documented in comprehensive surveys [1,2].

The texture classification problem is conventionally divided into the two subproblems of feature extraction and classification [1,2]. To improve the overall quality of texture classification, either the quality of the texture features or the quality of the classification algorithm must be improved. This paper focuses on the improvement of texture feature quality, extending earlier preliminary work published in [3] and the work in [4].

There has been longstanding interest in developing robust features for texture classification with strong invariance to rotation, illumination changes, view point variations, perspective projection changes, nonrigid deformations and occlusions [5–12,16]. In other words, the major challenge is to develop texture features which not only are highly discriminative to inter-class textures, but are also robust to one or more intra-class variations. This paper focuses on the important problem of robust gray-scale and rotation invariant texture features.

Rotation invariant feature extraction is usually a complex process, with some steps treated with special care and being computationally demanding [17,18]. Our research is motivated by the concluding remark—"a very useful direction for future research is therefore the development of powerful texture measures that can be extracted and classified with a low computational complexity" in the recent excellent comparative study of Randen and Husøy [2]. Remarkable work along these lines is the LBP set of features [8], the filtering features of Schmid [19] and Leung and Malik [20], and the recent work of Varma and Zisserman [11] who showed that raw image pixel features from local image patches can outperform popular filter bank features such as the rotation-invariant MR8 features.

The dimensionality of patch features can cause severe limitations in the applicability of the patch method of Varma and Zisserman [11]. In order to circumvent this problem, Liu and Fieguth [4] introduced the use of random projections (RPs), a universal, information-preserving dimensionality-reduction technique, to project the patch vector space to a compressed patch

* Corresponding author. Tel.: +86 731 84573479x807; fax: +86 731 84518730.
E-mail addresses: dreamliu2010@gmail.com (L. Liu), pfieguth@uwaterloo.ca (P. Fieguth), dclausi@engmail.uwaterloo.ca (D. Clausi), kuangyeats@vip.sina.com (G. Kuang).

0031-3203/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2011.10.027

space without a loss of salient information, claiming that the performance achieved by random features can outperform patch features, MR8, and LBP features.

Even though impressive classification performance was obtained in [4] using RP features, the approach is sensitive to image rotation. Fig. 1 serves as a motivational example for the exploration of the proposed rotation invariant scheme, contrasting the distributions of sorted and unsorted random projections. A texture image produces a cluster in the random feature space, and rotating the texture causes the cluster to be spread along some curve in panels (a,c). Sorting the patches before taking a random projection (b,d) limits the extent to which the cluster is spread along a path, leading to an impressive improvement in class locality and separability. This proposed approach will be referred to as sorted random projection (SRP).

The proposed SRP classifier preserves all of the computational simplicity, universality, and high classification performance advantages of the basic RP classifier. We will show the SRP features to be robust, invariant to image rotation locally, yet very discriminative, allowing us to take advantage of the powerful BoW model [5,6,11] for global rotation invariant texture classification. Furthermore, our method avoids the careful design steps and expensive computational cost involved in some local feature descriptors such as RIFT [5,6], SPIN [5,6] and SIFT [6].

The rest of this paper is organized as follows. Section 2 reviews the background literature for rotation invariant texture classification. Sections 2 and 3 respectively review the RP classifier and develop the proposed SRP classifier. In Section 5, we evaluate the capabilities of the proposed features with extensive experiments on *seven* popular texture datasets, summarized in Table 1, and present comparisons with current state-of-the-art classifiers on each dataset.

2. Background

2.1. BoW and texture classification

There has been a great interest in using a "Bag of Words" (BoW) approach for texture classification [5,6,8,9,11,20,22]. The BoW model, representing texture images statistically as histograms over a universal texton dictionary learned from local features, has proven widely effective for texture classification. This BoW model encompasses two popular paradigms for texture classification, dense and sparse, summarized in Fig. 2.

The dense approach uses local features pixel by pixel over the image, requiring feature extraction, texton selection, image histogram learning and classification (see the upper arrows in Fig. 2). Noticeable work along these lines includes [8,9,11,20,22]. In contrast, the sparse approach uses local features at a sparse set of interest points, with a corresponding sequence of key point detection, feature extraction at key points, texton selection, signature representation of image and classification (lower arrows in Fig. 2) [5,6].

There are three reasons why the BoW approach is popular for invariant texture classification. First, the representation is built on powerful local texture feature descriptors, which can be made insensitive to local image perturbations such as rotation, affine changes and scale. Second, the use of histogram as the statistical characterization for each image is globally invariant to these same changes. Third, the representation can be compared using standard distance metrics, allowing robust classification methods such as support vector machines to be employed.

2.2. Rotation invariant texture features

The general approach to developing rotation invariant techniques has been to modify successful non-rotation invariant techniques,

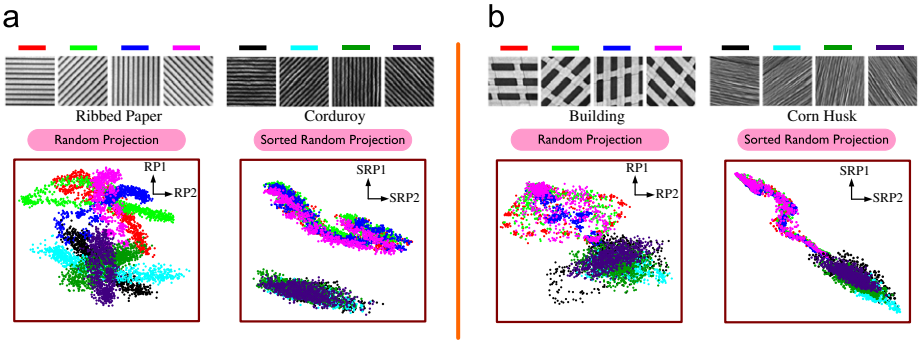

Fig. 1. Consider random projections of four different textures at varying orientations. The scatter plots in the bottom row show the random projections for a large number of extracted texture patches. Relative to random projections (a,c), it is clear that the sorted random projections in (b,d) offer superior class separability and compactness.

Table 1
Summary of texture datasets used in the experiments.

Texture dataset	Dataset notation	Image rotation	Controlled illumination	Scale variation	Significant viewpoint	Texture classes	Sample size	Samples per class	Samples in total
CUReT	$\mathcal{D}^{C}$	√	√			61	200 × 200	92	5612
Brodatz	$\mathcal{D}^{B}$					111	215 × 215	9	999
CUReTRot	$\mathcal{D}^{CRot}$	√	√			61	140 × 140	92	5612
BrodatzRot	$\mathcal{D}^{BRot}$	√				111	128 × 128	9	999
UIUC	$\mathcal{D}^{UIUC}$	√		√	√	25	640 × 480	40	1000
UMD	$\mathcal{D}^{UMD}$	√		√	√	25	320 × 240	40	1000
KTH-TIPS	$\mathcal{D}^{KT}$		√	√		10	200 × 200	81	810

such as rotation invariance derived from gray level cooccurrence histograms by averaging the histograms over different angles. As features based on Markov Random Field (MRF) models are inherently rotation-variant, several schemes have been introduced to obtain rotation invariance such as the Circular Simultaneous Auto-Regressive model [21], the MultiResolution Simultaneous AutoRegressive model by Mao and Jain [24], and the Anisotropic Circular Gaussian MRF model by Deng and Clausi [25]. Similarly there are filter-based approaches [2,41], such as circular Gabor filters, rotation invariant Daubechies wavelet transform features [38], steerable filters [39], and the log-polar transform [40], among others [2,41].

Although several factors can contribute to invariance, the invariance of local features plays a key role. However, there are limitations with many of the existing local features for rotation invariant classification. Existing outstanding local features, such as SIFT, RIFT, GLOH and SPIN [5,6], in general only work well on a sparse set of points, and they often need to be combined with each other to achieve satisfying performance. Furthermore, an important factor affecting the performance of sparse features is image resolution, since key-point extraction tends to not work well on low resolution images [6].

Similarly for dense features, such as LBP, the joint responses of various features [9,20], though being rotation invariant, have been shown to give unsatisfactory classification performance [4,11]. Patch features [11] and random features [4], which are at the state-of-the-art in texture classification, are not rotation invariant.

2.3. Random projection

Random projection (RP) [4,26,27,42,43] refers to the technique of projecting a set of points from a high-dimensional space to a randomly chosen low-dimensional subspace. Such techniques have been used in a wide variety of applications, including information retrieval, face recognition [28] and machine learning [29,30]. They represent a computationally simple and efficient means of preserving texture structure without introducing significant distortion.

The information-preserving and dimensionality reduction power of RP is firmly demonstrated by the theory of compressed sensing (CS) [31–33], which states that for sparse and compressible signals, a small number of nonadaptive linear measurements in the form of random projections can capture most of the salient information in the signal. Moreover, RP also provides a feasible solution to the well-known Johnson–Lindenstrauss (JL) lemma [27], which states that a point set in a high-dimensional Euclidean space can be mapped down onto a much smaller space, having a dimensionality logarithmic in the number of points, but with the distances between the points approximately preserved.

3. A review of the RP classifier

The RP classifier proposed by Liu and Fieguth [4] is divided into three stages. Readers are referred to [4] for details of the development of the basic RP classifier. Each image $\mathbf{I}$ is first represented as a collection of patch vectors, denoted as $\{\underline{\boldsymbol{x}}\}$. Some number of RP measurements

$$\underline{\boldsymbol{y}} = \boldsymbol{\Phi}\underline{\boldsymbol{x}} \tag{1}$$

are computed for each patch vector, where the entries of the projection matrix $\boldsymbol{\Phi}$ are independent, zero mean, unit variance normal. The texture classification process is as follows:

1. Compressed texton dictionary learning: The compressed patch vectors $\{\underline{\boldsymbol{y}}_i\}$ are extracted from the training images, then are normalized by Weber's Law [4,9,11]
$$\underline{\boldsymbol{y}} \leftarrow \underline{\boldsymbol{y}}\left[\frac{\log(1+\|\underline{\boldsymbol{y}}\|_2/0.03)}{\|\underline{\boldsymbol{y}}\|_2}\right] \tag{2}$$
to further enhance intensity invariance. The normalized compressed patch vectors from each texture class are clustered using k-means. The resulting K cluster centers $\{\underline{\boldsymbol{w}}_i\}$ from all texture classes form a universal compressed texton dictionary of size KC, where C is the number of texture classes considered.
2. Histogram learning: A histogram $\underline{\boldsymbol{h}}$ of compressed textons is learned for each particular training sample by labeling each of its pixels with the closest texton. Next, the histogram of texton labelings is normalized to define a representation vector for the image. Each texture class then is represented by a set of histogram models corresponding to the training samples of that class.
3. Classification: Given an image to classify, its associated normalized texton histogram is learned in the same way as during training, allowing the texture to be classified using a nearest neighbor classifier in histogram space, where the distance between two histograms is measured using a χ^2 statistic
$$\chi^2(\underline{\boldsymbol{h}}_1,\underline{\boldsymbol{h}}_2) = \frac{1}{2}\sum_{k=1}^{CK}\frac{[\underline{\boldsymbol{h}}_1(k)-\underline{\boldsymbol{h}}_2(k)]^2}{\underline{\boldsymbol{h}}_1(k)+\underline{\boldsymbol{h}}_2(k)} \tag{3}$$

4. The proposed SRP classifier

The RP classifier uses random measurements of local image patches to perform texture classification, however the fact that the image patch features are not rotationally invariant can be a serious limitation. Existing general methodologies to achieve rotation invariance in the patch vector representation include three main approaches:

1. Add randomly rotated versions of the training samples to the training set when learning textons. This results in clusters having many more points and a much greater spread (Fig. 1), clearly posing storage and processing challenges, and also creating challenges in clustering the texton space, since the required number of cluster centers K increases with cluster spread.
2. Estimate the dominant gradient orientation of the local patch and align the patch with respect to it [5,6,9,11]. The dominant orientation estimates tend to be unreliable, especially for blob regions which lack strong edges at the center, and corner regions which have more than one dominant orientation. In addition to reliability, finding the dominant orientation for each local patch is computationally expensive.
3. Marginalize the intensities weighted by the orientation distribution over angle, or compute multilevel histograms at fixed distances from the center of a patch (e.g. the SPIN descriptor adopted in [5,6]).

Based on the above discussion, and again motivated by the striking classification results by Liu and Fieguth [4], we would like to further capitalize on the RP approach by proposing a robust variant. The most attractive result in [4] is that the method is conceptually simple, untuned, computationally straightforward, and the dimensionality requirements are quite modest—a small number of simple random projections of a local patch contains enough information for successful texture classification. We therefore introduce a strategy to achieve rotation invariance and, at the same time, retain the simplicity, effectiveness and the theoretical support of the RP classifier [4], avoiding the disadvantages of existing rotation invariant approaches.

We begin by revisiting the intuitive motivational example in Fig. 1. An ensemble of patches extracted from a texture produces

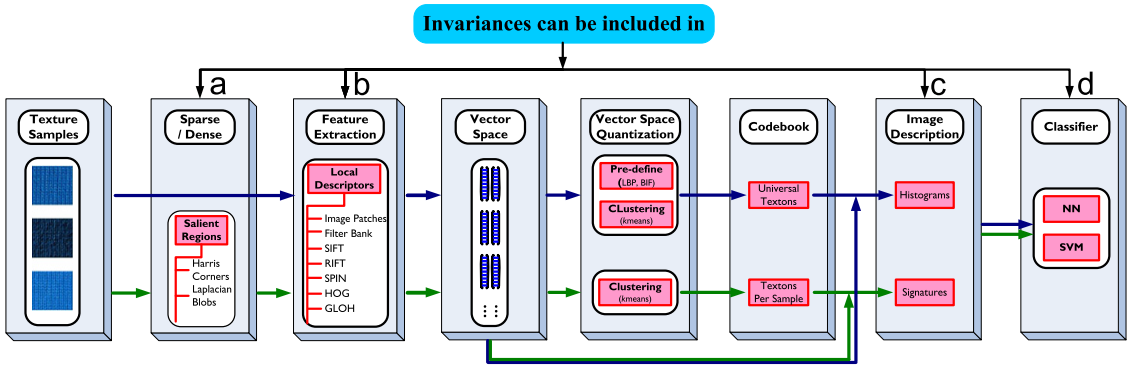

Fig. 2. A review of the commonly used components of the two popular BoW-based frameworks for texture classification: the dense approach (top arrows) and the sparse approach (bottom arrows). Texture classifiers obeying some invariance generally require careful attention to (a,b) local invariant feature extraction, (c) global invariant image representation, and (d) invariances obtained by the classifier, such as invariant similarity measures. This paper concentrates on invariant features (b).

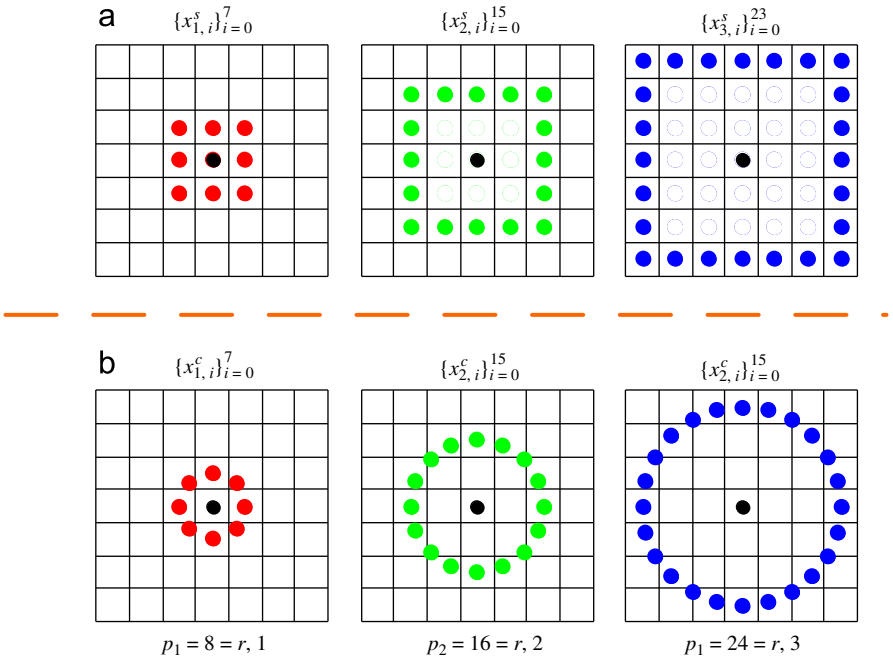

Fig. 3. Examples of symmetric neighborhood structures in three sizes: (a) square symmetric neighborhoods, (b) circular symmetric neighborhoods. The black solid circle is the reference pixel of the neighborhood.

a cluster of points in some feature space. Rotating the texture thus causes the cluster to be spread along some curve, as can be observed in the figure. The more spread and complex a feature point class is, the more difficult the corresponding clustering encounters, since the clustering then requires many more code-vectors (i.e., cluster centers or textons) to cover the feature subspace that the texture class occupies, leading to an increased computational burden.

Based on the above analysis, we want to re-localize the cluster and make it more compact. To this end, we replace a local texture patch vector with a sorted one. Essentially, instead of using compressed patch vector $\underline{\boldsymbol{y}} = \boldsymbol{\Phi}\underline{\boldsymbol{x}}$ for classification, we just use

$$\underline{\boldsymbol{y}} = \boldsymbol{\Phi} sort(\underline{\boldsymbol{x}}) \tag{4}$$

where we sort over all (or parts) of $\underline{\boldsymbol{x}}$. Fig. 1(b) and (d) is a preliminary exploration of the idea, showing a scatter plot of the joint output of a pair of random projections of the sorted patch vectors (i.e., SRP features) sampled from the images. In contrast to the scatter plots of the RP features shown in Fig. 1(a) and (c), the clusters corresponding to different rotated images overlap, becoming one cluster, while also being easily distinguishable from the other classes, i.e., class separability is preserved or even improved.

4.1. Sorted pixel values

We are now in a position to present the proposed approach. Texture is characterized not only by the gray value at a given pixel, but also by the gray value "pattern" in a neighborhood surrounding the pixel. To start with, consider the $(2a+1)\times(2a+1)$ square neighborhood of Fig. 3(a), where we modify the RP classifier by replacing the RP measurements of local patch vectors with the RP measurements of sorted local patch vectors

$$\underline{\boldsymbol{y}} = \boldsymbol{\Phi}\underline{\boldsymbol{x}}^{\mathrm{Glob}} \tag{5}$$

$$\underline{\boldsymbol{x}}^{\mathrm{Glob}} = [x_{0,0}, sort([x^s_{1,0}, \ldots, x^s_{a,p_a-1}])]^T \tag{6}$$

$x_{0,0}$ corresponds to the value of the center pixel of the local patch, $\{x^s_{r,i}\}_i$, $0 \le i < p_r, 1 \le r \le a$ correspond to the values of the rth concentric square of pixels, as illustrated in Fig. 6(a), such that all of the pixels in the square neighborhood around the center pixel are sorted. Since sorting ignores the ordering of elements in $\underline{\boldsymbol{x}}$, $\underline{\boldsymbol{x}}^{\mathrm{Glob}}$ is clearly rotation invariant (excepting the effects of pixellation).

Fig. 4 validates the basic behavior of the SRP features. The classification results show a jump (from below 80% to above 90%) in classification performance, compared to basic RP in classifying the textures in the challenging UIUC texture database. Given the strength of the RP results in [4], this surprising result firmly confirms the effectiveness of the sorting strategy.

Clearly, global sorting provides a poor discriminative ability, since crudely sorting over the whole patch (the center pixel excluded) not only discards all orientation information, but also leads to an ambiguity of the relationship among pixels from different scales. Clearly a natural extension of global sorting is to only sort pixels of the same circumference. We propose two kinds of multiscale sorting schemes: the rectangular or square neighborhood and the circular neighborhood, both shown in Fig. 3. We will refer to texture classification based on these two multiscale sorting schemes (illustrated in Fig. 6(b) and (c) respectively) as SRP Square and SRP Circular.

As illustrated in Fig. 6(b), the SRP Square feature is

$$\underline{\boldsymbol{y}} = \boldsymbol{\Phi}\underline{\boldsymbol{x}}^{\mathrm{Sqr}} \tag{7}$$

$$\underline{\boldsymbol{x}}^{\mathrm{Sqr}} = [x_{0,0}, sort([x^s_{1,0}, \ldots, x^s_{1,p_1-1}]), \ldots, sort([x^s_{a,0}, \ldots, x^s_{a,p_a-1}])]^T \tag{8}$$

The sorting is applied over each square ring, with p_r pixels in ring r, with r determining the spatial scale of the operator. Since a square is rotation invariant only to angles which are multiples of $\pi/2$, we clearly expect the performance of this approach to be limited.

Similarly, as shown in Fig. 6(c), we propose a SRP Circular feature

$$\underline{\boldsymbol{y}} = \boldsymbol{\Phi}\underline{\boldsymbol{x}}^{\mathrm{Circ}} \tag{9}$$

$$\underline{\boldsymbol{x}}^{\mathrm{Circ}} = [x_{0,0}, sort([x^c_{1,0}, \ldots, x^c_{1,p_1-1}]), \ldots, sort([x^c_{a,0}, \ldots, x^c_{a,p_a-1}])]^T \tag{10}$$

where $\{x^c_{r,i}\}_{i=0}^{p_r-1}$ denotes the values of the p_r neighbors of $x_{0,0}$ on a circle of radius r. Fig. 3(b) illustrates such circularly symmetric neighbor sets. Those locations which do not fall exactly on a pixel are estimated by interpolation. To maintain consistency with the square neighbor sets, the values of (p_r, r) are chosen to be (8,1), (16,2), (24,3), etc.

In our proposed approach, sorting provides stability against rotation, while sorting at each circumference individually preserves the overall between-circumference spatial information. In this way, a compromise is achieved between the conflicting requirements of greater geometric invariance on the one hand and greater discriminative power on the other. As can be seen from Fig. 5, sorting over concentric squares or circular rings both offer an improvement over global sorting.

For illustration purposes, we give a specific example to show that the proposed SRP descriptor is resistant to rotational perturbations. Fig. 7 shows a texture image at eight orientations, with corresponding texton histograms. As can be seen, the texton distributions change very little for the SRP method, despite the fact that there is considerable intra-class variation due to image rotation.

Relationship of sorting to SPIN: Our circularly sorted pixel value descriptor is closely related to the SPIN descriptor. The SPIN descriptor, proposed by Lazebnik et al. [5], is a rotation-invariant multilevel histogram of intensities within an image region, computed by dividing the image region into concentric rings of equal width and concatenating all the sub-histograms, each of which is simply the histogram of the intensity values of pixels located at a concentric ring. The difficulty of the SPIN approach is that the histogram requires gray-value quantization, causing a loss of information and resulting in ambiguities, as many different values are mapped to the same histogram entry. Furthermore, because the traditional histogram is a discontinuous mapping and has bad clustering properties, Lazebnik et al. [5] implemented the SPIN histogram as a soft histogram, which increased the computational complexity significantly. On this basis we believe that the proposed sorting scheme is expected to give better performance than SPIN.

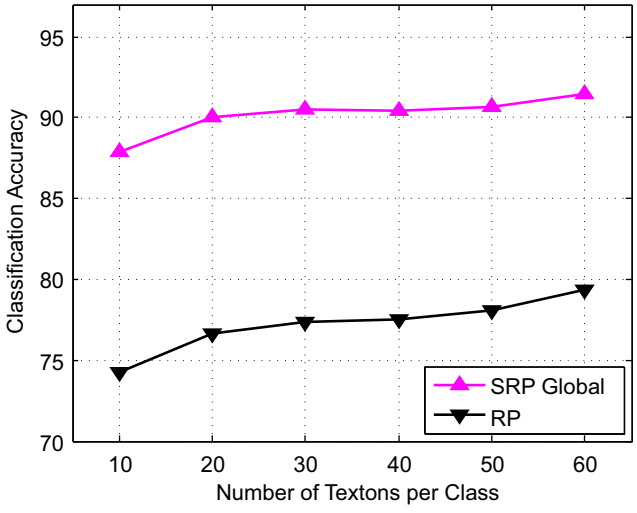

Fig. 4. Comparison of a primitive sorted-RP versus RP: classification accuracy as a function of the number of textons per class on $\mathcal{D}^{\mathrm{UIUC}}$, given 20 samples per class for training, using a patch of size 9×9, and a dimensionality of 30. Results are averaged over tens of random partitionings of the training and testing sets.

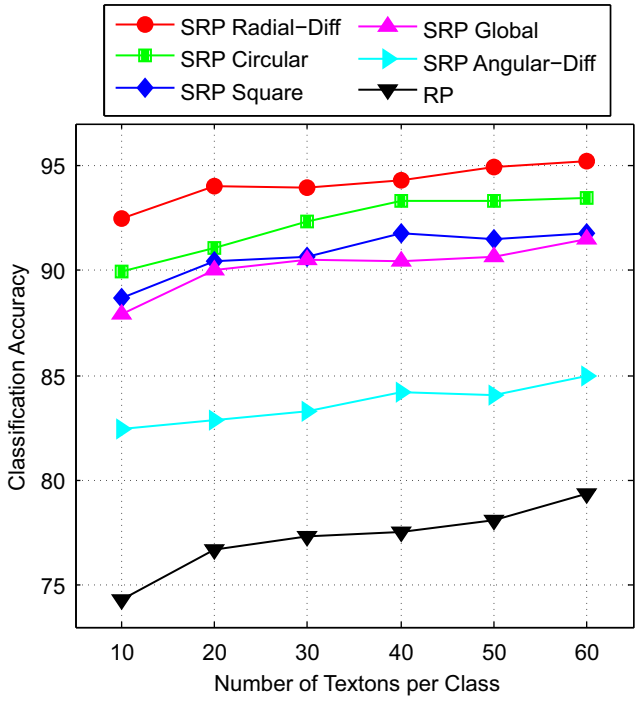

Fig. 5. Like Fig. 4, but comparing all of the proposed sorting schemes with basic RP.

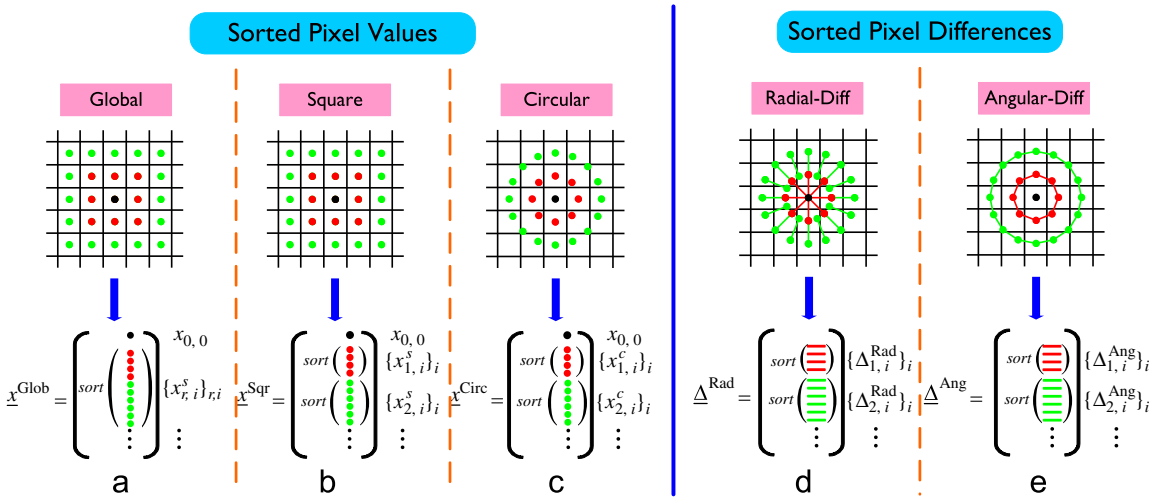

Fig. 6. Sorting schemes on a local image patch: sorting pixels (a–c) or sorting pixel differences (d, e). The pixels may be taken natively on a square grid (a, b) or interpolated to lie on rings of constant radius (c–e).

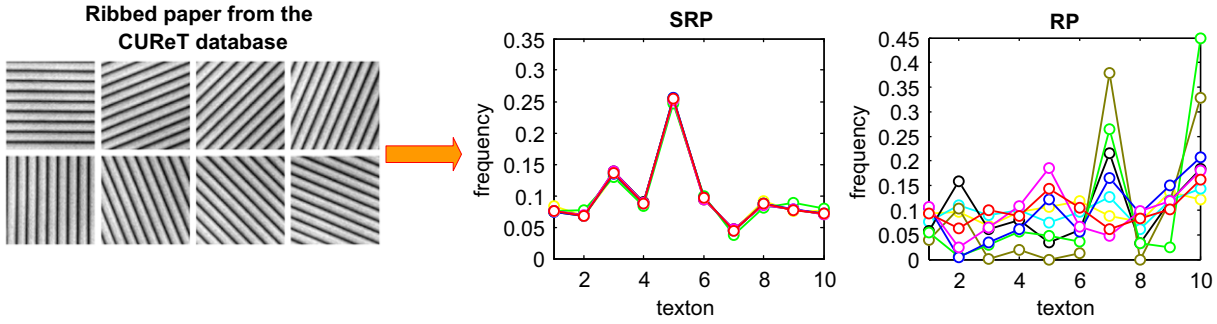

Fig. 7. *SRP circular versus basic RP*: The rotated texture images at eight uniformly sampled orientations are shown for "Ribbed Paper". We train 10 textons, and plot the texton histogram or distribution for each of the eight images. Clearly the histograms are nearly identical for the sorted random projections, but not at all for unsorted ones. Therefore, in the case of SRP, we can reasonably expect having one image present in the training set will allow the other rotated samples to be classified correctly. In contrast, for RP the learned models are quite dissimilar.

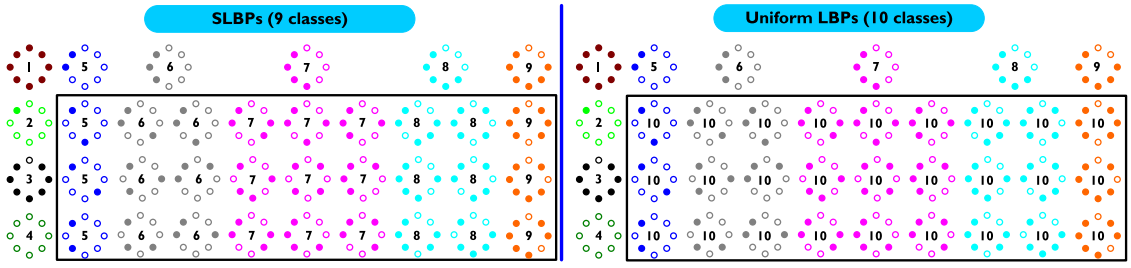

Fig. 8. Comparing the groups of the SLBPs and the uniform LBPs. The 36 unique rotation invariant binary patterns that can occur in the circularly symmetric neighbor set of (8,1) are grouped into 9 groups and 10 groups respectively. Solid and hollow circles correspond to bit values of 0 and 1 (see [8] for more details). The numbers inside the patterns correspond to their class membership.

Relationship of Sorted LBPs and Uniform LBPs: To illustrate the power of the proposed sorting schema, we could perform sorting on LBPs to obtain Sorted LBPs (SLBPs). With no loss of generality, suppose we take a patch size of 3×3. According to the study in [8], it was demonstrated that better classification performance on rotation texture classification can be achieved by grouping the rotation invariant LBPs (LBPROTs) into nine groups of rotation invariant uniform patterns (LBPRIUs), illustrated on the first row and the first column of the right panel (i.e., groups 1–9) of Fig. 8, and one group of all the other so-called "nonuniform" patterns. When applying our sorting operator on the LBPROTs, we can easily obtain nine groups of SLBPs, as illustrated in the left panel of Fig. 8. The pattern groups labeled 1, 2, 3, and 4 are exactly the same for SLBPs and LBPRIUs (see Fig. 8).

We can easily observe that the difference is that the regroup strategy for the 27 nonuniform LBPROTs (shown in square area of Fig. 8): for the LBPRIUs, all the 27 rotation invariant patterns are grouped into another single class, denoted as the "miscellaneous". Note that nothing would prevent us from regrouping the other 27 rotation invariant patterns, so for our SLBPs, the other 27 rotation invariant patterns are distributed into several SLBPs.

4.2. Sorted pixel differences

Sorting each ring of pixels loses any sense of spatial coupling, whereas textures clearly possess a great many spatial relationships. Therefore we propose sorting radial or angular differences, illustrated in Fig. 6(d) and (e). It is worth noting that the significance of

pairwise gray-level information for texture discrimination was advocated in 1962 by Julesz [34]. Since then, gray level cooccurrence histograms, gray level difference histograms, and features derived from them have been successfully used in a large number of texture analysis studies [8,35,36].

We propose signed pixel differences in radial and angular directions on a circular grid. Clearly the sorted radial pixel differences are also robust to illumination changes, since the signed differences are not affected by changes in mean luminance. Conceptually, our approach may be traced back to early papers on the extraction of local difference histograms [35,36], where the emphasis is placed on the pairwise pixel relationships. Later on, codebook based quantization of signed differences of neighboring pixels was presented in [8], where the obvious advantage of signed gray-level differences over absolute differences was verified quantitatively in texture classification.

However, the differences used in [8] are computed in a very small 3×3 neighborhood and LBPs focus on the differences between the center pixel and its neighborhood only, in contrast to the proposed radial and angular pixel differences over larger patches. In particular, radial differences encode the between-circumference structure, thus sorted radial differences can achieve rotation invariance while preserving the relationship between pixels of different rings. Formally, the sorted radial and angular difference descriptors are computed as

$$\underline{\Delta}^{\mathrm{Rad}} = [sort(\Delta^{\mathrm{Rad}}_{1,0},\ldots,\Delta^{\mathrm{Rad}}_{1,p_1-1}),\ldots,sort(\Delta^{\mathrm{Rad}}_{a,0},\ldots,\Delta^{\mathrm{Rad}}_{a,p_a-1})]^T \tag{11}$$

$$\underline{\Delta}^{\mathrm{Ang}} = [sort(\Delta^{\mathrm{Ang}}_{1,0},\ldots,\Delta^{\mathrm{Ang}}_{1,p_1-1}),\ldots,sort(\Delta^{\mathrm{Ang}}_{a,0},\ldots,\Delta^{\mathrm{Ang}}_{a,p_a-1})]^T \tag{12}$$

where

$$\Delta^{\mathrm{Rad}}_{r,i} = x^c_{r,i} - x^c_{r-1,i*p_{r-1}/p_r}, \quad \Delta^{\mathrm{Ang}}_{r,i} = x^c_{r,i} - x^c_{r,i-1} \tag{13}$$

Then the SRP Radial-Diff Classifier and the SRP Angular-Diff Classifier are based on

$$\underline{\boldsymbol{y}} = \boldsymbol{\Phi}\underline{\Delta} \tag{14}$$

Fig. 5 plots the classification results comparing RP and all of the proposed SRP features on $\mathcal{D}^{\mathrm{UIUC}}$. The results show that all of the SRP classifiers perform significantly better than the RP classifier, where sorted radial differences performed the best.

The motivation of the RP method in [4] was to avoid feature extraction. We do confess that our proposed sorting and differencing schemes may seem a tiny bit like feature extraction, however we are deliberately choosing only the most basic and trivial operations (differencing, sorting). We avoid trying a wide variety or more complex features, and are promoting only very simple invariants.

5. Experimental evaluation

To demonstrate the effectiveness of the proposed approach for rotation invariant texture classification, we have used *seven* different datasets, summarized in Table 1 and described in detail in the following sections, derived from the five most commonly used texture sources: CUReT [9,11], Brodatz [37], UIUC [5], UMD [16], and KTH-TIPS [7]. For comparative evaluation, the proposed methods have been compared with six state-of-the-art texture classification approaches (detailed in Section 5.1). The presentation of the experimental results is divided into the following three parts.

Experiment #1 is carried out with the purposes of measuring the performance of the proposed methods on the Brodatz database $\mathcal{D}^{\mathrm{B}}$ and the CUReT database $\mathcal{D}^{\mathrm{C}}$, both of which have no significant rotation variations, and of comparing the performance of the proposed SRP approach with the RP classifier [4] on the same datasets as Liu and Fieguth [4].

Experiment #2 is conducted to test the classification performance of the proposed methods for rotation invariant texture classification, involving a more challenging setup with randomly rotated textures generated synthetically from datasets $\mathcal{D}^{\mathrm{B}}$ and $\mathcal{D}^{\mathrm{C}}$, resulting in two new texture datasets $\mathcal{D}^{\mathrm{BRot}}$ and $\mathcal{D}^{\mathrm{CRot}}$.

Experiment #3 involves three challenging datasets ($\mathcal{D}^{\mathrm{UIUC}}$, $\mathcal{D}^{\mathrm{UMD}}$ and $\mathcal{D}^{\mathrm{KT}}$) with texture images which have large intraclass variations such as significant viewpoint changes, significant scale variation, or even nonrigid surface deformations other than rotation, to show that the proposed approaches can also be generalized to successfully classify very challenging databases and achieve very high performance.

Implementation details: To make the comparisons as meaningful as possible, we use the same experimental settings as [4,11]. Each sample is intensity normalized to have zero mean and unit standard deviation. All results are reported over 50 random partitions of training and testing sets. Each extracted SRP vector is normalized via Weber's law. Histograms/χ^2 and nearest neighbor classifier (NNC) are used. Half of the samples per class are randomly selected for training and the remaining half for testing, except for $\mathcal{D}^{\mathrm{B}}$ and $\mathcal{D}^{\mathrm{BRot}}$, where three samples are randomly selected as training and the remaining six as testing.

5.1. Methods in the comparison study

Our specific experimental goal is to compare the proposed SRP approach with the current state-of-the-art, which we identify to be the following:

RP [4]: Each local patch of size $\sqrt{n}\times\sqrt{n}$ is reordered into an n-dimensional patch vector, and is then compressed into a $m<n$-dimensional vector using random projections. Both training and testing are performed in the compressed domain.

Patch [11]: Each local patch of size $\sqrt{n}\times\sqrt{n}$ is reordered into an n-dimensional patch vector. Both training and testing are performed in the patch domain.

LBP [8]: The rotationally invariant, uniform LBP texton dictionary at different scales, $\mathrm{LBP}^{\mathrm{riu2}}_{8,1}$, $\mathrm{LBP}^{\mathrm{riu2}}_{8,1+16,2}$, $\mathrm{LBP}^{\mathrm{riu2}}_{8,1+16,2+24,3}$, $\mathrm{LBP}^{\mathrm{riu2}}_{8,1+16,2+24,3+24,4}$, $\mathrm{LBP}^{\mathrm{riu2}}_{8,1+16,2+24,3+24,4+24,5}$ advocated in [44,45]. For simplicity, in the remainder of this paper these LBP textons are denoted as 1-scale, …, 5-scale respectively.

The method of Lazebnik et al. [5]: The basic idea is to first characterize the texture by two types of elliptic regions: Harris-affine corners and Laplacian-affine blobs. The regions are then normalized to circles such that the local descriptor is invariant to affine transform. Two descriptors (SPIN and RIFT) are used for feature extraction from each region. The signature/EMD framework is used.

The method of Zhang et al. [6]: Based on the method of Lazebnik et al. [5], such that Zhang et al. use three types of descriptors (SPIN, RIFT and SIFT) and a kernel SVM classifier; the EMD SVM classifier is very computationally demanding.

WMFS [16]: Based on a combination of wavelet transform and multifractal analysis of scale-normalized texture images, where the scales are estimated from local affine invariant blob areas [5].

5.2. Image data and experimental setup

The *Brodatz* dataset ($\mathcal{D}^{\mathrm{B}}$) [37] has 111 texture classes with only one sample of size 640×640 in each class. Performing classification on the entire database is challenging due to the relatively large number of texture classes, the small number of examples for each class, and the lack of intra-class variation. Furthermore, some inhomogeneous textures are hardly from the same class when divided into non-overlapping subimages, as shown in Fig. 9. Some textures appear very similar up to a scale change, which is clearly stated in the original source [37]. We use the

same dataset as Liu and Fieguth [4], Lazebnik et al. [5] and Zhang et al. [6], dividing each texture image into nine 215×215 non-overlapping subimages.

For the *CUReT* dataset ($\mathcal{D}^{C}$) [9,11] we use the same subset of images as [4,9,11]. These images are captured under different illuminations with seven different viewing directions. Intra-class variations due to lighting geometry is very severe for this database. The effects of specularities, interreflections, shadowing, and other surface normal variations are plainly evident and can be seen in the left subfigure of Fig. 10.

The *BrodatzRot* and *CUReTRot* datasets ($\mathcal{D}^{BRot}$, $\mathcal{D}^{CRot}$) are synthesized by randomly rotating each of the database images by a random angle, uniformly between 0° and 360°. To avoid boundary problems when the images were rotated, only the center portions of the images were used.

The *UIUC* dataset ($\mathcal{D}^{UIUC}$) [5] has been designed to require local invariance. Textures are acquired under significant scale and viewpoint changes, arbitrary rotations, and uncontrolled illumination conditions. Furthermore, the database includes textures with varying local affine deformations, and nonrigid deformation of the imaged texture surfaces. The large intra-class variations can be observed from some example textures, plotted in Fig. 11(a), making the database very challenging for classification.

The *UMD* dataset ($\mathcal{D}^{UMD}$) [16] has been designed in a similar way as $\mathcal{D}^{UIUC}$ by the authors in [16], who realized that the images in $\mathcal{D}^{UIUC}$ were not large enough for their proposed feature vector computation. The database consists of high resolution images of size 1280×960, with arbitrary rotations, significant viewpoint changes and scale differences present. Compared with $\mathcal{D}^{UIUC}$, the database does not have textures with nonrigid deformation. In this paper, our experiments were performed on a lower resolution version of the original UMD database, with each image down-sampled to 240×320 using bilinear interpolation.

The *KTH-TIPS* dataset ($\mathcal{D}^{KT}$) [7] contains 10 texture classes with each class having 81 images, captured at nine lighting and rotation setups and nine different scales spanning two octaves (relative scale changes from 0.5 to 2). Example images with scale changes are shown in Fig. 12. $\mathcal{D}^{KT}$ extends CUReT by imaging new samples of 10 of the CUReT textures over a range of scales. Although KTH-TIPS is designed that it is possible to

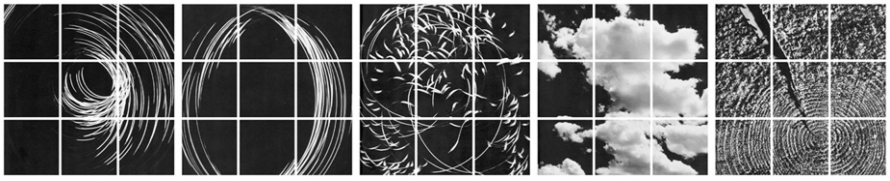

Fig. 9. Five nonhomogeneous textures (D43, D44, D45, D91 and D97) from the Brodatz database.

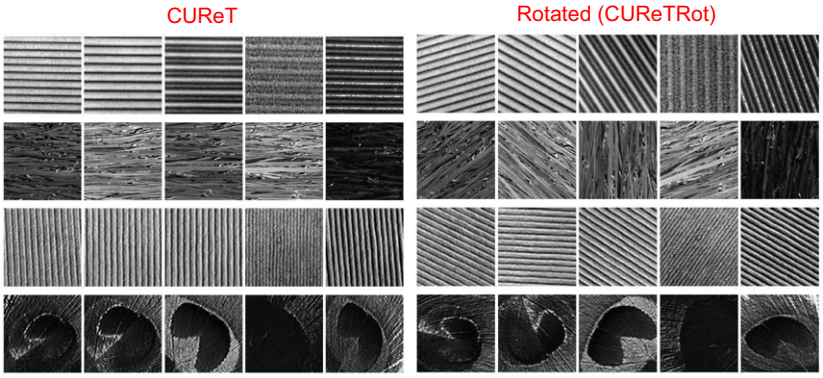

Fig. 10. The left subfigure shows five samples each from material numbers 38, 40, 42 and 57 from the CUReT [4] dataset. The right panel shows the synthetically rotated versions (i.e., from the CUReTRot dataset) of the same samples according to a randomly selected angle.

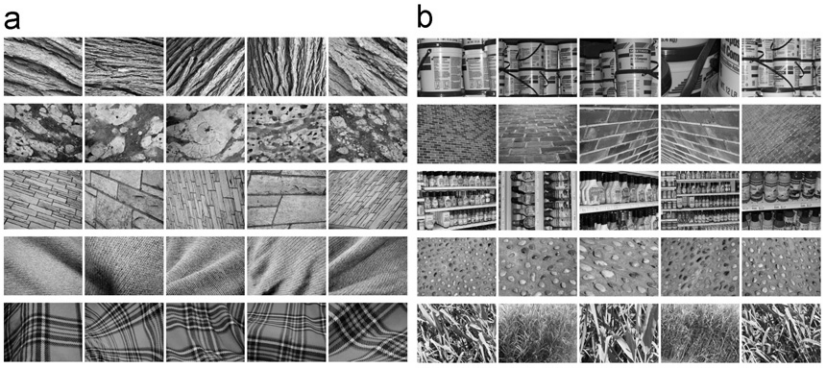

Fig. 11. Samples from five texture classes in (a) the UIUC dataset and (b) the UMD dataset, showing the large intra-class changes, including random rotation, large viewpoint variation, and large scale changes.

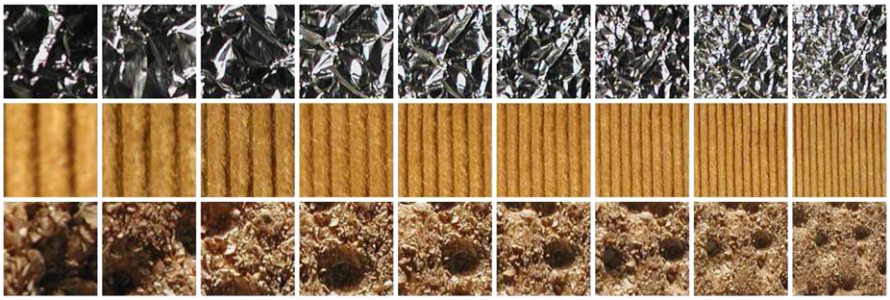

Fig. 12. Example textures from the KTH-TIPS database, showing samples of nine scales.

combine it with CUReT in testing, we follow Zhang et al. [6] in treating it as a stand-alone dataset.

Remarks. Both $\mathcal{D}^{\mathrm{UIUC}}$ and $\mathcal{D}^{\mathrm{UMD}}$ represent a major improvement over $\mathcal{D}^{\mathrm{C}}$ in that materials are imaged under significant viewpoint variations and some in $\mathcal{D}^{\mathrm{UIUC}}$ also have considerable surface deformations. A drawback is that these are also much smaller than $\mathcal{D}^{\mathrm{C}}$, both in the number of classes as well as the number of images per class. Variations in lighting geometry in $\mathcal{D}^{\mathrm{UIUC}}$ and $\mathcal{D}^{\mathrm{UMD}}$ are less severe than those for $\mathcal{D}^{\mathrm{C}}$ and $\mathcal{D}^{\mathrm{KT}}$. Nevertheless, as far as scale and other viewpoint variations are concerned, the $\mathcal{D}^{\mathrm{UIUC}}$ and $\mathcal{D}^{\mathrm{UMD}}$ databases are by far the most challenging and we therefore test the proposed features on them.

5.3. Experimental results

5.3.1. Experiment #1

The results on $\mathcal{D}^{\mathrm{B}}$ are shown in Fig. 13(a) and Table 2. We can observe that the improvement of the proposed SRP Radial-Diff approach over the basic RP classifier is quite significant. Even compared with the state-of-the-art methods of Lazebnik et al. [5] and Zhang et al. [6], our SRP Radial-Diff approach performs the best, and is significantly superior to that of the method of Lazebnik et al. [5] which uses a nearest neighbor classifier.

In terms of $\mathcal{D}^{\mathrm{C}}$, since Liu and Fieguth [4] have conducted extensive experimental evaluation and comparisons to other methods on this dataset, here we only provide results (shown in Table 3) for comparing with RP. The performance penalty in Table 3 for incorporating rotation invariance is very modest. In contrast, the method of Lazebnik et al. using a nearest neighbor classification achieves only $72.5\% \pm 0.7\%$. This result is somewhat surprising but highlights the inadequacies of sparse, affine adaptation based methods. Even when multiple high dimensional descriptors are combined with multiple detectors and sophisticated SVMs employed, i.e., the method of Zhang et al., the sparse method results improve to only $95.3\% \pm 0.4\%$ [6].

5.3.2. Experiment #2

Table 4 presents results for datasets $\mathcal{D}^{\mathrm{CRot}}$ and $\mathcal{D}^{\mathrm{BRot}}$. From these tables, we can see that all of the SRP classifiers significantly outperformed both the RP classifier [4] and the Patch method [11], and with the RP classifier outperforming the Patch method. These results confirm the solid power of the proposed methods for rotation and illumination invariant texture classification.

Again, we can observe that our SRP Radial-Diff approach consistently outperformed all the other methods in comparison: around 20% and 30% higher than both the RP [4] and Patch [11] methods, and around 4% and 9% higher than LBP [8] for $\mathcal{D}^{\mathrm{CRot}}$ and $\mathcal{D}^{\mathrm{BRot}}$ respectively. The excellent results for our proposed approaches demonstrate their suitability for rotation invariant texture classification.

5.3.3. Experiment #3

The results for dataset $\mathcal{D}^{\mathrm{UIUC}}$, the same dataset used by Lazebnik et al. [5] and Zhang et al. [6], are shown in Table 5 and Fig. 13(b). As expected, among the RP methods, SRP Radial-Diff clearly outperformed all other methods. Furthermore the radial-difference method significantly outperforms LBP, outperforms the method of Lazebnik et al. [5] using a nearest neighbor classification, and performs only slightly below the method of Zhang et al. [6]. This latter result should be interpreted in light of the fact that Zhang et al. use scale invariant and affine invariant channels and a more advanced classifier (EMD kernel SVMs), which is important for $\mathcal{D}^{\mathrm{UIUC}}$ where some textures have significant scale changes and affine variations. Moreover, note that some textures in $\mathcal{D}^{\mathrm{UIUC}}$, which are imaged at acute angles, tend to have sections that are out of focus. This affects the results: failure to recognize the same texture from two quite different viewpoints does not necessarily indicate a failure to generalize between viewpoints or rotations.

In the case of dataset $\mathcal{D}^{\mathrm{UMD}}$, we also did some experiments on the original high resolution dataset $\mathcal{D}^{\mathrm{UMD}}$. Image patches of size 9×9, 11×11 and 13×13 are tried while using 60 textons per class, leading to classification rates are $98.22\% \pm 0.31\%$, $98.72\% \pm 0.18\%$ and $98.41\% \pm 0.21\%$ respectively. We found that the classification accuracy was almost unaffected by image resolution, therefore in most of our experiments we used 240×320 images.

We compare our UMD results to two state-of-the-art methods—the RP classifier [4] and the WMFS approach of Xu et al. [16], shown in Table 5 and Fig. 13(c). It is seen that both SRP Square and SRP Radial-Diff can outperform the method of Xu et al. What is curious is that our rotation invariant SRP classifier achieves such a good performance (99.13%, almost perfect) despite the large scale changes (as illustrated in Fig. 11(b)) present in the UMD dataset. In contrast to the results on $\mathcal{D}^{\mathrm{UIUC}}$, the results obtained for the other methods are also relatively high, indicating the relatively low level of difficulty of the dataset $\mathcal{D}^{\mathrm{UMD}}$.

Table 6 lists the results for dataset $\mathcal{D}^{\mathrm{KT}}$. Note that $\mathcal{D}^{\mathrm{KT}}$ has controlled imaging and a small number of texture classes (only 10 classes). Textures in this dataset have no obvious rotation, though they do have controlled scale variations. From Table 6, we can see that SRP Radial-Diff again performs the best, outperforming all methods in the extensive comparative survey of Zhang et al. [6].

All the experimental results with a simple NNC classifier presented previously have firmly demonstrated the effectiveness and robustness of the proposed SRP features. Motivated by the work in [7,6,13,16], where a more sophisticated support vector machines (SVMs) classifier was used, we further examine the performance of the proposed SRP features with a SVMs classifier. We use $K(\underline{\boldsymbol{h}}_i, \underline{\boldsymbol{h}}_j) = \exp(-\gamma \chi^2(\underline{\boldsymbol{h}}_i, \underline{\boldsymbol{h}}_j))$. The kernel parameter γ for SVMs is found by cross-validation within the training set. The values of the parameters and of SVMs are specified using a grid search scheme. In this work, the publicly available *LibSVM*

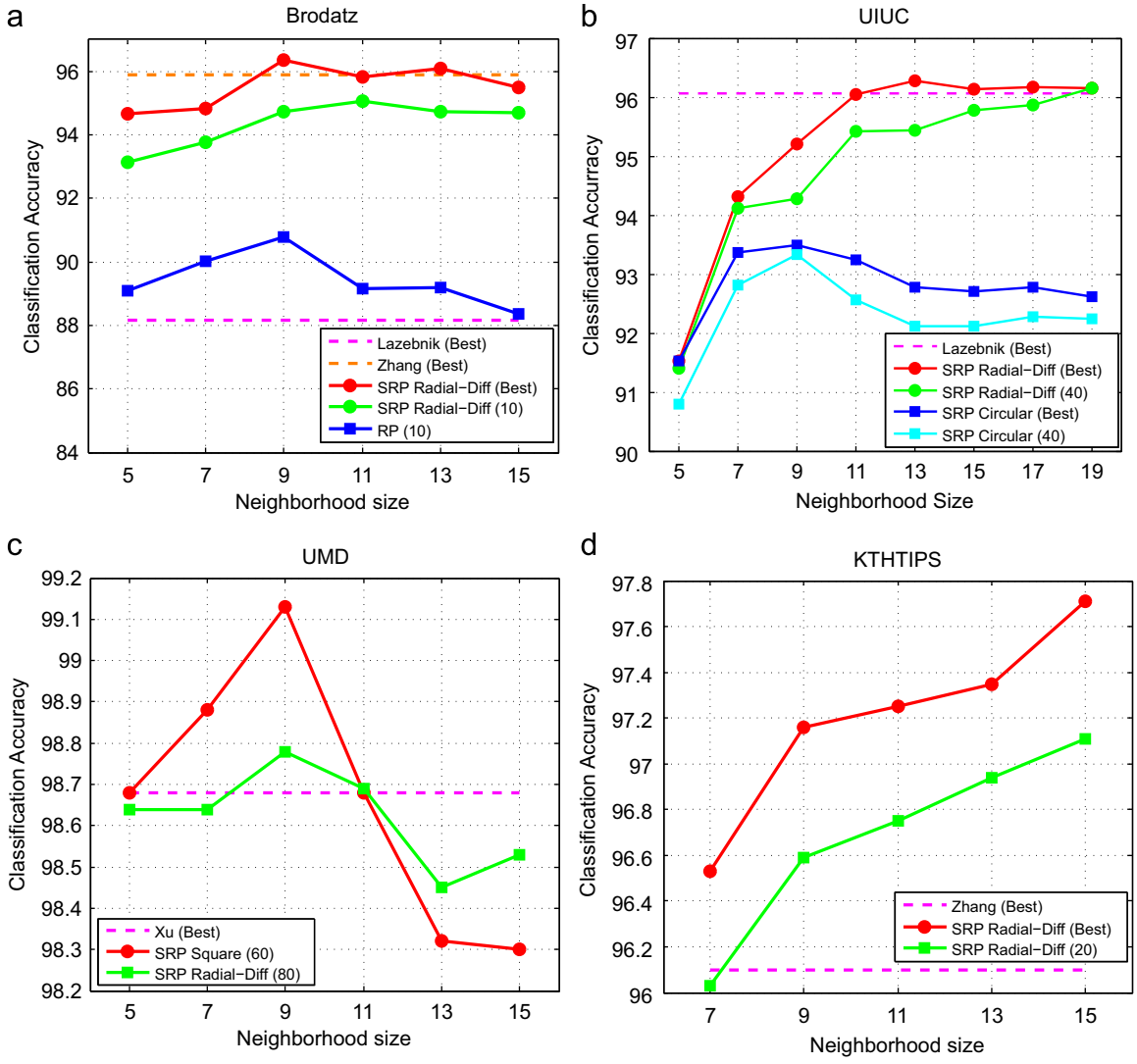

Fig. 13. Comparison of classification performance of various methods: (a) results on the $\mathcal{D}^{B}$. The SRP Radial-Diff Best curve shows the best results obtained by varying K up to 40. (b) Results on $\mathcal{D}^{UIUC}$. The SRP Radial-Diff Best curve shows the best results obtained by varying K up to 80. Similarly, SRP Circular Best curve is obtained by varying K up to 60. (c) Results on $\mathcal{D}^{UMD}$. (d) Results on $\mathcal{D}^{KT}$ and the SRP Radial-Diff Best curve shows the best results obtained by varying K up to 40. The results for "Lazebnik (Best)", "Zhang (Best)" and "Xu (Best)" are the highest classification accuracies, directly quoted from their original papers [5,6,16] respectively. The bracketed numbers denote the number of textons K used per class.

Table 2
Experimental results for $\mathcal{D}^{B}$: all results for our proposed approach are obtained by number of textons used per class $K=10$, except SRP Radial-Diff (Best), which are the best obtained by varying K up to 40. Results of Lazebnik et al. and Zhang et al. are quoted directly from [5,6].

Method	Patch size				
	5×5	9×9	11×11	13×13	15×15
Dimensionality	10	30	40	50	60
RP	$89.10\%\pm1.41\%$	$90.78\%\pm1.26\%$	$89.17\%\pm0.93\%$	$89.19\%\pm1.00\%$	$88.35\%\pm1.47\%$
SRP Global	$83.56\%\pm1.66\%$	$81.89\%\pm1.64\%$	$81.98\%\pm1.51\%$	$80.27\%\pm1.59\%$	$79.80\%\pm1.81\%$
SRP Square	$84.03\%\pm1.70\%$	$87.16\%\pm0.65\%$	$85.77\%\pm1.49\%$	$86.37\%\pm1.57\%$	$86.75\%\pm1.60\%$
SRP Circular	$86.70\%\pm1.23\%$	$88.38\%\pm1.06\%$	$87.36\%\pm1.00\%$	$86.09\%\pm1.25\%$	$85.68\%\pm1.46\%$
SRP Radial-Diff	$93.13\%\pm1.15\%$	$94.74\%\pm0.91\%$	$95.05\%\pm0.76\%$	$94.73\%\pm1.08\%$	$94.68\%\pm1.43\%$
SRP Radial-Diff (Best)	$94.67\%\pm1.15\%$	$\mathbf{96.34\%\pm0.38\%}$	$95.83\%\pm0.66\%$	$96.10\%\pm0.53\%$	$95.48\%\pm1.43\%$
SRP Angular-Diff	$88.91\%\pm1.51\%$	$89.79\%\pm1.31\%$	$92.37\%\pm1.57\%$	$90.80\%\pm1.07\%$	$90.54\%\pm1.16\%$
Scale	1	2	3	4	5
LBP	$80.93\%\pm1.27\%$	$87.52\%\pm0.88\%$	$88.87\%\pm1.62\%$	$89.69\%\pm1.65\%$	$89.94\%\pm0.53\%$
Lazebnik best [5]	88.15%				
Best from Zhang [6]	95.9%				

library [46] is employed. The parameters C and γ are searched exponentially in the ranges of $[2^{-5},2^{18}]$ and $[2^{-15},2^{8}]$, respectively, with a step size of 2^{1} to probe the highest classification rate.

Table 7 gives a comprehensive summary of the results for our proposed approach with 12 recent state-of-the-art results. We can observe that our approach scores very well across all five

commonly used datasets, producing what we believe to be the best reported result on the CUReT, Brodatz, KTH-TIPS, and UMD datasets, and for UIUC database our classification rate 98.4% is very close to the best reported results (98.9%). The best score 99.22% on CUReT is reported by Broadhurst [14], who used a Gaussian Bayes classifier with marginal filter distributions. Noticeably, the method CG-MSBIF proposed by Crosier and Griffin [22], which uses a multiscale metric to boost classification performance, also performs fairly good. However, CG-MSBIF produces a rather long histogram feature vector (1296×8) for each texture image, which is memory-demanding and time-consuming in the classification stage.

At last, Table 8 summarizes the number of methods compared and the number of texture databases tested in their original work for various state-of-the-art methods. To the best of our knowledge, only our work has shown extensive experimental results on all five challenging databases, whereas most papers in the field test approaches on one or two texture databases, or even just part of one database. It is precisely the scope and breadth of our results that allow us to claim robustness and universality. We point out further that we have compared our approach to a large range of approaches (13 approaches from the literature). Considering our broad testing of texture databases and the extensive comparison with the state-of-the-art, we believe our approach has significant impact and merit.

Table 3
Comparisons of classification results of the Basic RP, the SRP Circular, and the SRP Radial-Diff on $\mathcal{D}^{C}$ with $K=10$.

Patch size	RP (%)	SRP Circular (%)	SRP Radial-Diff (%)
7×7	96.80	96.37	96.33
9×9	96.91	96.58	96.61

6. Conclusions

In this paper we have presented a novel, near feature-extraction free approach for rotation invariant texture classification. We proposed a simple but surprisingly effective sorting scheme to make the local patch vector invariant to image rotation locally. The texture classification is then performed based on the compressed random projection features of the sorted patches. We developed two more

Table 4
Experimental results for the SRP classifier on $\mathcal{D}^{CRot}$ and $\mathcal{D}^{BRot}$, all results are obtained by ourselves with $K=10$ except for those labeled (Best) which are obtained by varying $K=10, 20, 30, 40$.

Method	$\mathcal{D}^{CRot}$					$\mathcal{D}^{BRot}$				
	Patch size					Patch size				
	7×7	9×9	11×11	15×15	19×19	5×5	7×7	9×9	11×11	15×15
Dimensionality	20	30	40	60	80	10	25	30	40	60
RP	74.12%	74.91%	75.31%	75.85%	75.42%	59.16%	57.08%	54.35%	52.63%	48.26%
SRP Global	93.21%	91.19%	91.65%	91.17%	89.33%	73.90%	74.62%	72.97%	72.22%	69.44%
SRP Square	94.37%	94.14%	93.82%	93.61%	93.37%	76.93%	77.72%	76.75%	75.27%	71.30%
SRP Circular	95.20%	95.22%	94.85%	94.38%	95.21%	76.55%	79.94%	79.45%	78.96%	74.74%
SRP Circular (Best)	95.66%	**96.53**%	96.31%	96.45%	95.55%	N/A	N/A	N/A	N/A	N/A
SRP Radial-Diff	93.81%	94.55%	94.74%	94.76%	95.01%	84.53%	**87.78**%	87.24%	87.65%	85.91%
SRP Radial-Diff (Best)	94.33%	95.01%	95.57%	95.76%	**96.11**%	85.53%	**88.63**%	88.16%	88.04%	87.22%
SRP Angular-Diff	86.09%	88.59%	89.18%	90.04%	89.36%	73.95%	77.76%	79.13%	79.73%	79.71%
Dimensionality	49	81	121	225	361	10	25	81	121	225
Patch	73.31%	74.47%	74.79%	75.15%	74.84%	57.93%	56.07%	52.98%	51.49%	47.95%
Scale	1	2	3	4	5	1	2	3	4	5
LBP^{riu2}	64.42%	84.23%	88.73%	91.09%	92.53%	60.45%	73.25%	76.69%	78.51%	79.43%

Table 5
Experimental results for $\mathcal{D}^{UIUC}$ and $\mathcal{D}^{UMD}$: all results are obtained by number of textons used per class $K=40$, except for those labeled (Best), which are the best obtained by varying K up to 80. Results of Lazebnik et al. [5], Zhang et al. [6] and Xu et al. [16] are quoted directly from the original papers.

Method	$\mathcal{D}^{UIUC}$					$\mathcal{D}^{UMD}$				
	Patch size					Patch size				
	5×5	9×9	11×11	13×13	15×15	7×7	9×9	11×11	13×13	15×15
Dimensionality	10	30	40	50	60	20	30	40	50	60
RP	79.60%	77.53%	77.56%	76.34%	76.30%	96.07%	96.23%	95.69%	94.83%	94.79%
SRP Global	90.56%	90.44%	91.23%	90.87%	91.53%	98.55%	98.64%	98.46%	98.54%	98.15%
SRP Square	N/A	N/A	N/A	N/A	N/A	98.73%	98.65%	98.46%	98.28%	98.30%
SRP Square (Best)	90.83%	91.77%	91.53%	91.59%	91.81%	98.88%	**99.13**%	98.68%	98.32%	98.39%
SRP Circular	90.80%	93.33%	92.56%	92.11%	92.11%	98.54%	98.43%	98.34%	97.75%	97.85%
SRP Radial-Diff	91.40%	94.28%	95.42%	95.43%	95.77%	98.48%	98.60%	98.19%	98.26%	98.27%
SRP Radial-Diff (Best)	91.52%	95.20%	96.05%	**96.27**%	96.13%	98.64%	98.78%	98.69%	98.45%	98.53%
SRP Angular-Diff	77.13%	84.19%	84.94%	86.46%	86.50%	96.69%	97.43%	97.20%	97.48%	97.33%
Scale	1	2	3	4	5	1	2	3	4	5
LBP	58.08%	75.64%	81.45%	84.10%	86.10%	92.96%	95.53%	95.35%	95.86%	95.65%
Other methods in comparison	Best from Lazebnik [5] 96.1% Best from Zhang [6] **98.7**%					WMFS best [16] 98.68%				

Table 6
Experimental results for $\mathcal{D}^{KT}$: all results for our proposed approach are obtained by $K=20$ except SRP Radial-Diff (Best), which are the best obtained by varying $K=10$, 20, 30, 40. Results of Zhang et al. are quoted directly from [6].

Method	Patch size				
	7×7	9×9	11×11	13×13	15×15
Dimensionality	25	30	40	50	60
RP	$95.61\%\pm0.65\%$	$95.17\%\pm0.61\%$	$94.56\%\pm0.36$	$94.05\%\pm0.53\%$	$94.20\%\pm0.67\%$
SRP Global	$93.93\%\pm0.76\%$	$93.33\%\pm0.85\%$	$93.67\%\pm0.91\%$	$92.57\%\pm0.96\%$	$92.88\%\pm0.82\%$
SRP Square	$96.12\%\pm0.59\%$	$94.74\%\pm0.78\%$	$95.40\%\pm0.78\%$	$95.52\%\pm0.66\%$	$95.58\%\pm0.60\%$
SRP Circular	$95.09\%\pm0.77\%$	$95.40\%\pm0.60\%$	$95.07\%\pm0.39\%$	$94.71\%\pm1.11\%$	$94.02\%\pm0.96\%$
SRP Radial-Diff	$96.03\%\pm0.89\%$	$96.59\%\pm0.53\%$	$96.75\%\pm0.64\%$	$96.94\%\pm0.52\%$	$\mathbf{97.11\%\pm0.56\%}$
SRP Radial-Diff (Best)	$96.53\%\pm0.30\%$	$97.16\%\pm0.36\%$	$97.25\%\pm0.44\%$	$97.35\%\pm0.30\%$	$\mathbf{97.71\%\pm0.49\%}$
SRP Angular-Diff	$90.99\%\pm1.46\%$	$91.07\%\pm0.67\%$	$90.87\%\pm1.09\%$	$91.87\%\pm1.24\%$	$90.81\%\pm0.62\%$
Best from Zhang [6]	$96.1\%\pm1.1\%$				

Table 7
Comparing the best classification scores achieved by the proposed SRP Radial-Diff using both a χ^2 kernel SVMs classifier and NNC classifier with the best reported classification scores for various methods on five datasets. The patch sizes used are 13×13, 9×9, 13×13, 9×9 and 9×9 for $\mathcal{D}^{C}$, $\mathcal{D}^{B}$, $\mathcal{D}^{KT}$, $\mathcal{D}^{UIUC}$, $\mathcal{D}^{UMD}$ respectively. The number of textons K used per class is 10, 10, 40, 40, and 40 respectively. Scores are as originally reported, except for those marked (∗) which are taken from the comparative study in Zhang et al. [6].

Methods	Databases				
	$\mathcal{D}^{C}$ 46 training samples per class	$\mathcal{D}^{B}$ three training samples per class	$\mathcal{D}^{KT}$ 41 training samples per class	$\mathcal{D}^{UIUC}$ 20 training samples per class	$\mathcal{D}^{UMD}$ 20 training samples per class
1. Our method (SVMs)	$\mathbf{99.05\%\pm0.28\%}$	$\mathbf{96.78\%\pm0.59\%}$	$\mathbf{99.11\%\pm0.36\%}$	$\mathbf{98.4\%\pm\%0.59}$	$\mathbf{99.15\%\pm0.36\%}$
2. Our method (NNC)	$98.52\%\pm0.19\%$	$\mathbf{96.34\%\pm0.38\%}$	$97.71\%\pm0.49\%$	$96.27\%\pm0.44\%$	$\mathbf{99.13\%\pm0.16\%}$
3. VZ-MR8 [9]	97.43%				
4. VZ-Patch [11]	98.03%	$92.9\%\pm0.8\%$(∗)	$92.4\%\pm2.1\%$ (∗)	$97.83\%\pm0.66\%$	
5. Hayman et al. [7]	$98.46\%\pm0.09\%$	$95.0\%\pm0.8\%$ (∗)	$94.8\%\pm1.2\%$ (∗)	$92.0\%\pm0.1.3\%$ (∗)	
6. Lazebnik et al. [5]	$72.5\%\pm0.7\%$ (∗)	88.15%	$91.3\%\pm1.4\%$ (∗)	96.03%	
7. Mellor et al. [15]		89.71 %			
8. Zhang et al. [6]	$95.3\%\pm0.4\%$	$95.9\%\pm0.6\%$	$96.1\%\pm1.2\%$	$\mathbf{98.7\%\pm0.4\%}$	
9. Brodhurst [14]	$\mathbf{99.22\%\pm0.34\%}$				
10. Varma and Ray [10]				$\mathbf{98.9\%\pm0.68\%}$	
11. CG-MSBIF [22]	$98.6\%\pm0.2\%$		$98.5\%\pm0.7\%$	$\mathbf{98.8\%\pm0.5\%}$	
12. Xu-MFS [12]				92.74%	93.93%
13. Xu-OTF [13]				97.40%	98.49%
14. Xu-WMFS [16]				**98.60%**	98.68%

Table 8
Summary: N_{cmp} denotes the number of originally compared methods. N_{db} denotes the number of texture databases originally tested.

Methods	Ours	VZ-MR8 [9]	VZ-Patch [11]	Hayman et al. [7]	Lazebnik et al. [5]
N_{cmp}	13	3	5	1	1
N_{db}	7	1	3	2	2
Methods	Mellor et al. [15]	Zhang et al. [6]	Brodhurst [14]	Varma and Ray [10]	
N_{cmp}	1	4	2	3	
N_{db}	2	4	1	2	
Methods	CG-MSBIF [22]	Xu-MFS [12]	Xu-OTF [13]	Xu-WMFS [16]	
N_{cmp}	7	2	3	4	
N_{db}	3	2	2	2	

invariants based on sorting the local pixel differences, either radially or angularly. The feature extraction process involves applying simple differencing and sorting to obtain rotation invariant features. This feature extraction process is quite efficient in contrast to the existing popular local feature descriptors.

The power and robust performance of our proposed approach have been demonstrated by classifying all the textures present in *seven* benchmark texture databases derived from five databases: the CUReT, Brodatz, UIUC, UMD and KTH-TIPS. The proposed method has been shown to match or surpass the state-of-the-art in texture classification, but with significant reductions in time and storage complexity. Among the sorted descriptors evaluated in this paper, the sorted radial difference descriptor is simple, yet it yields excellent performance across all databases. The proposed SRP approach outperforms all known classifiers on the Brodatz, rotated Brodatz, rotated CUReT, UMD and KTH-TIPS databases.

Acknowledgments

The authors would like to thank the School of Electronic Science and Engineering of National University of Defense

Technology, the Department of Systems Design and Engineering of the University of Waterloo, and the China Scholarship Council.

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] T. Randen, J. Husøy, Filtering for texture classification: a comparative study, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (4) (1999) 291–310.

[3] L. Liu, P. Fieguth, G. Kuang, Compressed sensing for robust texture classification, in: ACCV, Oral Presentation, 2010.

[4] L. Liu, P. Fieguth, Texture classification from random features, IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (3) (2012). doi:10.1109/TPAMI.2011.145.

[5] S. Lazebnik, C. Schmid, J. Ponce, A sparse texture representation using local affine regions, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (8) (2005) 1265–1278.

[6] J. Zhang, M. Marszalek, S. Lazebnik, C. Schmid, Local features and kernels for classification of texture and object categories: a comprehensive study, International Journal of Computer Vision 73 (2) (2007) 213–238.

[7] E. Hayman, B. Caputo, M. Fritz, J.-O. Eklundh, On the significance of real-world conditions for material classification, in: ECCV, vol. 4, 2004, pp. 253–266.

[8] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (7) (2002) 971–987.

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

[10] M. Varma, R. Garg, Locally invariant fractal features for statistical texture classification, in: ICCV, 2007, pp. 1–8.

[11] M. Varma, A. Zisserman, A statistical approach to material classification using image patches, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (11) (2009) 2032–2047.

[12] Y. Xu, H. Ji, C. Fermüller, Viewpoint invariant texture description using fractal analysis, International Journal of Computer Vision 83 (1) (2009) 85–100.

[13] Y. Xu, S. Huang, H. Ji, C. Fermüller, Combining powerful local and global statistics for texture description, in: CVPR, 2009.

[14] R. Broadhurst, Statistical estimation of histogram variation for texture classification, in: Proceedings of the Fourth International Workshop on Texture Analysis and Synthesis, 2005, pp. 25–30.

[15] M. Mellor, B.W. Hong, M. Brady, Locally rotation, contrast, and scale invariant descriptors for texture analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence 30 (1) (2008) 52–61.

[16] Y. Xu, X. Yang, H. Ling, H. Ji, A new texture descriptor using multifractal analysis in multi-orientation wavelet pyramid, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010, pp. 161–168.

[17] T.N. Tan, Scale and rotation invariant texture classification, in: IEE Colloquium Texture Classification: Theory and Applications, 1994, pp. 3/1–3/3.

[18] J. Zhang, T. Tan, Brief review of invariant texture analysis methods, Pattern Recognition 35 (3) (2002) 735–747.

[19] C. Schmid, Constructing models for content-based image retrieval. in: CVPR, vol. 2, 2001, pp. 39–45.

[20] T. 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.

[21] R.L. Kashyap, A. Khotanzad, A model-based method for rotation invariant texture classification, IEEE Transactions on Pattern Analysis and Machine Intelligence 8 (4) (1986) 472–481.

[22] M. Crosier, L.D. Griffin, Using basic image features for texture classification, International Journal of Computer Vision 88 (3) (2010) 447–460.

[24] J. Mao, A.K. Jain, Texture classification and segmentation using multiresolution simultaneous autoregressive models, Pattern Recognition 25 (2) (1992) 173–188.

[25] H. Deng, D.A. Clausi, Gaussian MRF rotation-invariant features for image classification, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (7) (2004) 951–955.

[26] S. Dasgupta, Experiments with random projections, in: Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence, 2000, pp. 143–151.

[27] S. Dasgupta, A. Gupta, An elementary proof of a theorem of Johnson and Lindenstrauss, Random Structures and Algorithms 22 (1) (2003) 60–65.

[28] J. Wright, A. Yang, A. Ganesh, S.S. Sastry, Y. Ma, Robust face recognition via sparse representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (2) (2009) 210–217.

[29] E. Bingham, H. Mannila, Random projection in dimensionality reduction: applications to image and text data, in: Proceedings of the Seventh ACM International Conference on Knowledge Discovery and Data Mining, 2001, pp. 245–250.

[30] X.Z. Fern, C.E. Brodley, Random projection for high dimensional data clustering: a cluster ensemble approach, in: Proceedings of the Twentieth International Conference on Machine Learning, 2003, pp. 186–193.

[31] E.J. Candès, T. Tao, Decoding by linear programming, IEEE Transactions on Information Theory 51 (12) (2005) 4203–4215.

[32] E. Candès, T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies? IEEE Transactions on Information Theory 52 (12) (2006) 5406–5425.

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

[34] B. Julesz, Visual pattern discrimination, Proceedings of the IRE 8 (2) (1962) 84–92.

[35] R.W. Conners, C.A. Harlow, A theoretical comparison of texture algorithms, IEEE Transactions on Pattern Analysis and Machine Intelligence 2 (3) (1980) 204–222.

[36] M. Unser, Sum and difference histograms for texture classification, IEEE Transactions on Pattern Analysis and Machine Intelligence 8 (1) (1986) 204–222.

[37] P. Brodatz, Textures: A Photographic Album for Artists and Designers, Dover, New York, 1966.

[38] R. Porter, N. Canagarajah, Robust rotation-invariant texture classification: wavelet, Gabor filter and GMRF based schemes, IEE Proceedings of Vision, Image, Signal Processing 144 (3) (1997) 180–188.

[39] W.T. Freeman, E.H. Adelson, The design and use of steerable filters, IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (9) (1991) 891–906.

[40] C-M. Pun, M.-C. Lee, Log-polar wavelet energy signatures for rotation and scale invariant texture classification, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (5) (2003) 590–603.

[41] S.C. Zhu, Y. Wu, D. Mumford, Filters, random fields and maximum entropy (FRAME): towards a unified theory for texture modeling, International Journal on Computer Vision 27 (2) (1998) 107–126.

[42] W.B. Johnson, J. Lindenstrauss, Extensions of Lipschitz mappings into a Hilbert space, in: Conference in Modern Analysis and Probability, 1984, pp. 189–206.

[43] D. Achlioptas, Database-friendly random projections, in: Proceedings of the Twentieth ACM Symposium on Principles of Database Systems, 2001, pp. 274–281.

[44] M. Pietikäinen, T. Nurmela, T. Mäenpää, M. Turtinen, View-based recognition of real-world textures, Pattern Recognition 37 (2) (2004) 313–323.

[45] T. Mäenpää, M. Pietikäinen, Multi-scale binary patterns for texture analysis, in: SCIA, 2003.

[46] C.-C. Chang, C.-J. Lin, LIBSVM: A Library for Support Vector Machines, Software available at 〈http://www.csie.ntu.edu.tw/∼cjlin/libsvm〉, 2001.

Li Liu received the B.S. degree in communication engineering and the M.S. degree in remote sensing and geographic information system from the National University of Defense Technology, Changsha, China, in 2003 and 2005, respectively, where she is currently pursuing the Ph.D. degree. She is currently a Visiting Student at the University of Waterloo, Ontario, Canada. Her current research interests include computer vision, texture analysis, pattern recognition and image processing.

Paul W. Fieguth (S'87–M'96) received the B.A.Sc. degree from the University of Waterloo, Ontario, Canada, in 1991 and the Ph.D. degree from the Massachusetts Institute of Technology, Cambridge, in 1995, both degrees in electrical engineering. He joined the faculty at the University of Waterloo in 1996, where he is currently Professor in Systems Design Engineering. He has held visiting appointments at the University of Heidelberg in Germany, at INRIA/Sophia in France, at the Cambridge Research Laboratory in Boston, at Oxford University and the Rutherford Appleton Laboratory in England, and with postdoctoral positions in Computer Science at the University of Toronto and in Information and Decision Systems at MIT. His research interests include statistical signal and image processing, hierarchical algorithms, data fusion, and the interdisciplinary applications of such methods, particularly to remote sensing.

David Clausi (S'93–M'96–SM'03) received the B.A.Sc., M.A.Sc., and Ph.D. degrees from the University of Waterloo, Waterloo, ON, Canada, in 1990, 1992, and 1996, respectively, all in systems design engineering. After completing his doctorate, he worked in the medical imaging field at Mitra Imaging, Inc., Waterloo. He started his academic career in 1997 as an Assistant Professor in geomatics engineering with the University of Calgary, AB, Canada. In 1999, he returned to his alma mater and was awarded tenure and promotion to Associate Professor in 2003. He is an active Interdisciplinary and Multidisciplinary Researcher. He has an extensive publication record, publishing refereed journal and conference papers in the diverse fields of remote sensing, computer vision, algorithm design, and biomechanics. His primary research

interest is the automated interpretation of SAR sea-ice imagery, in support of operational activities of the Canadian Ice Service. The research results have successfully led to commercial implementations. Dr. Clausi has received numerous scholarships, conference paper awards, and two Teaching Excellence Awards.

Gangyao Kuang received the B.S. and M.S. degrees from the Central South University of Technology, Changsha, China, in 1988 and 1991, respectively, and the Ph.D. degree from the National University of Defense Technology, Changsha, in 1995. Since 1996, he has been the Codirector of the Remote Sensing Information Processing Laboratory, National University of Defense Technology, where he has worked on synthetic aperture radar (SAR) signal and image processing, automatic target detection and recognition, information fusion, and various remote sensing projects. He is currently a Professor in the School of Electronic Science and Engineering, National University of Defense Technology. He is the author/coauthor of over 200 papers and one book. His current interests include remote sensing, SAR image processing, change detection, SAR ground moving target indication, and the classification of polarimetric SAR images.