



**HAL**  
open science

## Texture Analysis with Shape Co-occurrence Patterns

Gang Liu, Gui-Song Xia, Wen Yang, Liangpei Zhang

► **To cite this version:**

Gang Liu, Gui-Song Xia, Wen Yang, Liangpei Zhang. Texture Analysis with Shape Co-occurrence Patterns. 2013. hal-00922002

**HAL Id: hal-00922002**

**<https://hal.science/hal-00922002>**

Submitted on 23 Dec 2013

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

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

# Texture Analysis with Shape Co-occurrence Patterns

Gang Liu, Gui-Song Xia  
 LIESMARS, Wuhan University,  
 Wuhan 430079, China

Email: {liugang.spl, guisong.xia}@whu.edu.cn

Wen Yang  
 EIS, Wuhan University,  
 Wuhan, 430079, China

Email: wenyang@whu.edu.cn

Liangpei Zhang  
 LIESMARS, Wuhan University,  
 Wuhan 430079, China

Email: zlp62@whu.edu.cn

**Abstract**—This paper presents a flexible shape-based texture method by investigating the co-occurrence patterns of shapes. More precisely, a texture image is represented by a tree of shapes, each of which is associated with several attributes. The modeling of texture is thus converted to characterize the tree of shapes. To this aim, we first learn a set of co-occurrence patterns of shapes from texture images, then establish a bag-of-words model on the learned shape co-occurrence patterns (SCOPs), and finally use the resulted SCOPs distributions as features for texture analysis. In contrast with existing work, the proposed method not only inherits the strong ability to depict geometrical aspects of textures and the high robustness to variations of imaging conditions from the shape-based texture method, but also provides a more flexible way to consider shape relationships and high-order statistics on the tree. To our knowledge, this is the first time to use co-occurrence patterns of explicit shapes as a tool for texture analysis. Experiments of texture retrieval and classification on various databases report state-of-the-art results and demonstrate the efficiency of the proposed method.

## I. INTRODUCTION

As a fundamental ingredient of image structures, texture conveys important cues in numerous processes of human visual perception. While, due to the high complexity of the structures in natural images, the modeling of texture is a challenging problem in image analysis and understanding. Over the past decades, tremendous investigations have been made in texture analysis, see e.g. [1]–[8], among which an active topic is developing texture models which can efficiently depict both the statistical and the geometrical aspects of textures and are robust to the variations of imaging condition as well.

In order to represent the structural aspects of textures, some mathematical tools, such as Gabor or wavelet-like analysis, are used to probe the atomic texture elements such as elongated blobs and terminators in images, and the marginal/joint distributions of the resulted responses are subsequently utilized to describe the statistical arrangement of texture [1], [2], [9]. The strong ability of such mathematical tools to handle multi-scale and oriented structures has made them one of the most popular tool for texture analysis. However, how to efficiently represent the highly geometrical aspects of textures, e.g. sharp transitions and elongated contours, is an open problem. To solve this problem, alternative wavelet-like approaches, e.g. Grouplet [10] and scattering transform [8], have been elaborated to enable more efficient representations of structured textures. In contrast with explicit models, patch-based method [11] provides another possibility for describing the structured aspects of textures, but it is not trivial to capture the multi-scale nature of textures by this kind of approaches.

Next, for meet the requirement of invariance with respect

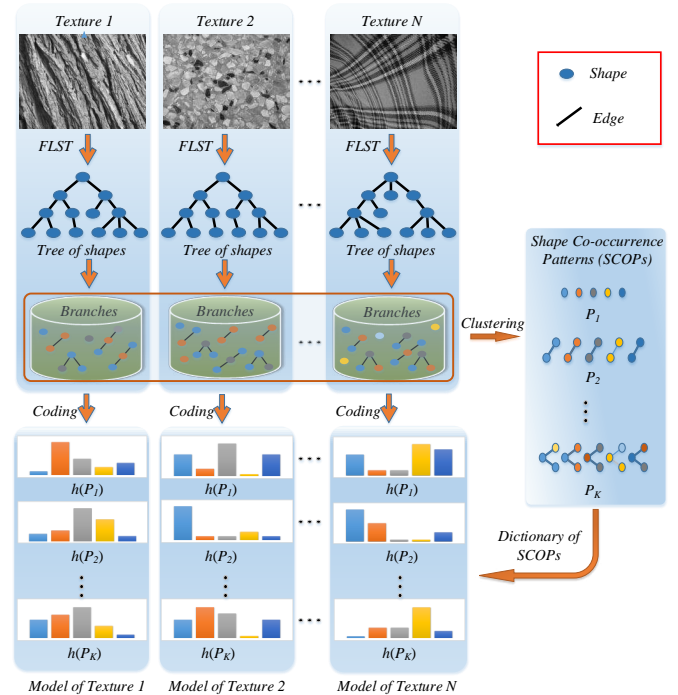


Fig. 1. The proposed texture analysis method with *shape co-occurrence patterns* (SCOPs). First images are represented by tree of shapes via Fast Level Set Transform (FLST) [12]. The branches of the trees are then collected and partitioned into different clusters, called SCOPs. An image is finally encoded by the learned SCOPs and the corresponding histograms are used as the texture model for analysis.

to viewpoint and illumination changes, many invariant texture descriptors have been proposed in the literature, including the rotation invariant local binary pattern (joint distribution of gray values on circular local neighborhoods) [13], the multi-fractal analysis method [14], etc.. Recently, several approaches rely on the extraction of local features that are individually invariant to some geometric transforms, such as scaling, rotating and shearing [3]. Compared with previous works dealing with invariant texture analysis, such locally invariant methods do not need any learning of the deformations. Alternatively, by relying on morphological operations, the shape-based invariant texture analysis (SITA) method [4] represents a texture by a tree of explicit shapes and shape attributes are locally normalized to achieve invariant texture description. SITA reported superior performance on invariant texture recognition. One limitation of this work, however, lies in its difficult to take into account high-order statistics of shapes, which has been reported as a crucial factor for texture discrimination [4].

Following the work of SITA [4], this paper contributes a more flexible shape-based texture analysis framework by investigating the co-occurrence patterns of shapes. The flowchart is illustrated in Figure 1. More precisely, given a texture, we first decompose it into a tree of shapes relying on a fast level set transformation [12], where each shape is associated with some attributes. We then learn a set of co-occurrence patterns of shapes from texture images, e.g. by K-means algorithm. Taking the learnt shape co-occurrence patterns as visual words, a bag-of-words model is finally established to describe the texture. In contrast with SITA [4], the proposed method provides a more flexible way to consider complex shape relationships and high-order statistics on the tree. Moreover, as we shall see, SITA can be regarded as a special case of the proposed one, where only marginal distributions and simple statistics of pair of shapes were taken into account. To our knowledge, the work in this paper is the first time to use explicit shape co-occurrence patterns for the analysis of textures. Several experiments on texture recognition demonstrate the efficiency of the proposed analysis method on various databases.

In the rest of the paper, we will briefly recall the background and related work in Section II. Next, in Section III we present the proposed SCOPs-based texture analysis method in detail. In Section IV, the ability of these features to classify and retrieve textures is demonstrated on various databases. Section IV concludes the paper with some remarks.

## II. BACKGROUND AND RELATED WORK

The proposed texture analysis method relying on shape co-occurrence patterns is inspired both by the shape-based texture analysis scheme and texture models using co-occurrence matrix, and is also closely related to the texton-based texture modeling paradigm. In what follows, we briefly recall these backgrounds.

1) *Co-occurrence patterns in textures*: Co-occurrence matrices [15], [16] is still a popular approach for texture analysis. It characterizes image textures with a set of sufficient non-parametric and low-order statistics of pixels. The preliminary use of co-occurrence matrices involved in statistics of pairwise pixel relationships in several predefined fashions. Local Binary Pattern (LBP) [17] extended the concept of co-occurrence by developing a framework for studying the statistics of co-occurrent binary patterns. Though, this kind of approaches demonstrated strong capability to characterize random and near random textures [17], one limitation of them is, however, a lack of consideration for large-scale and geometrical structures in texture images.

2) *Texton-based texture analysis paradigm*: Julesz [18] found that the first-order statistics of “a few local conspicuous features”, called *texton*, are significant for texture discrimination. The texton theory in fact led to a kind of structural approaches to texture description, which first probe texture primitives as local features and then investigate their organization. An implicit way to implement the texton theory is to use wavelet-like analysis, such as Gabor filter banks, to probe atomic texture elements in texture images and then utilize the resulted filtering responses to describe the underlying statistical texture features [1], [2], [8].

An alternative implementation is to explicitly detect atomic texture elements beforehand and model the spatial arrangements of such textons. For instance, Zhu *et al.* [9] detected textons by using a number of image bases with deformable spatial configurations, which are learned from static texture images. Lafarge *et al.* [19] first defined a set of geometric objects, e.g. segment, line, circle, band etc., as texture elements and then detected those elements and studied their organization through Markov Point Process (MPP). Note that a more general manner is to extract small image patches from texture, cluster them into textons and finally investigate the underlying statistics. Lazebnik *et al.* [3] extended this idea by removing the redundancy between patch-based textons through utilizing interest regions in images. Compared with the former implicit models, the explicit models can more easily handle structured parts in texture, such as edges and bars, which emerge in high-resolution image textures. However, the computation or detection of such textons is not trivial. It is also worth noticing that modeling the interactions between textons may involve heavy computation.

3) *Shape-based invariant texture analysis*: The shape-based invariant texture analysis scheme, called SITA in brief, is first proposed in [4]. It is based on a complete morphological image representation, called topographical map [20], which is made of all the connected components of the level lines of images. The idea is to decompose an images into a tree of shapes (by using fast level set transformation (FLST) [12]) and then develop texture features from the shape ensembles and their relationships. Note that SITA is also a texton-based texture analysis method, if we regard the explicit shapes to be textons in the paradigm. As the analysis scheme use explicit shapes to characterize textures, it demonstrated strong ability to depict the geometrical aspects of texture images, and thanks to the flexible normalization of geometrical transformations from individual shapes, it showed high efficiency for achieving geometric invariant texture features and reported superior performance on invariant texture recognition task. However, the scheme simply assumed that the shape attributes are independent and modeled textures by several marginal distributions. One limitation of such a scheme lies in its difficulty to handle the relationships of shapes, corresponding to high-order statistics on the tree, which is reported to be important for texture discriminations, see [4] for more details, where only the scale ratio between pair of shapes were considered and performed better than other attributes.

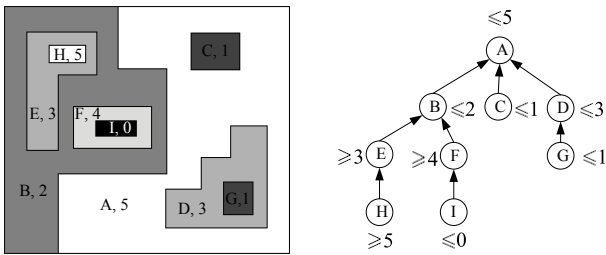
As mentioned before, this paper tries to combine the shape-based texture analysis scheme and the co-occurrence patterns methods in the texton-based paradigm. It can inherit, from the shape-based texture method, both the strong ability to model the geometrical aspects of textures and the high robustness to imaging conditions changes, but also provides a more flexible way to consider high-order statistics on the tree by investigating the co-occurrence patterns of shapes.

## III. TEXTURE MODELS WITH SCOPs

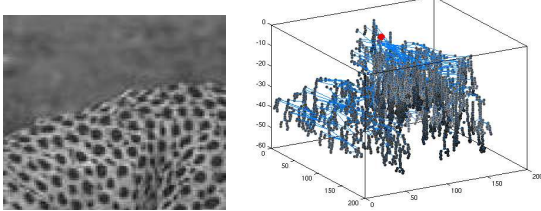
This section presents the proposed texture analysis method using shape co-occurrence patterns (SCOPs) in detail, following a short recall on the basis of SITA [4].

## A. Texture modeling by SITA

1) *Topographic map*: For a gray-scale image  $u$ , the upper and lower level sets are defined respectively as  $\chi_\lambda(u) = \{x \in \Omega; u(x) \geq \lambda\}$  and  $\chi^\lambda(u) = \{x \in \Omega; u(x) \leq \lambda\}$ , for  $\lambda \in R$ . The topographic map [20] of the image  $u$  is made of the connected components of the topological boundaries of the upper level sets of the image (they could be equivalently defined from the lower level sets). Observe that the connected components of upper level sets (respectively of the lower level sets) are naturally embedded in a tree structure. Monasse and Guichard combined these two redundant tree structures, by drawing on the notion of *shape*, and developed an efficient way to compute a hierarchical representation of images [12], named FLST<sup>1</sup> as mentioned before. A *shape* is defined as the interior of a level line (the boundary of a level set). Figure 2 shows examples of the topographic map representation of a synthetic image and a texture.



(a) Representation of a synthetic image by its topographic map. Left: the original digital image; Right: representation of the image by its tree of shapes, where  $(A, B, \dots, I)$  denote the corresponding shapes.



(b) Representation of a texture image by its topographic map.

Fig. 2. Topographic map representation, *i.e.* tree of shapes, of images.

2) *SITA* [4]: As the topographic map provides a complete representation of images, the modeling of texture  $u$  is converted to the modeling of the tree of shapes. The invariant texture features first rely on classical shape moments, then make use of the hierarchical structure of the topographic map.

The  $(p + q)$ th order central moments of shape  $s$  is defined as

$$\mu_{pq} = \int \int_s (x - \bar{x})^p (y - \bar{y})^q dx dy, \quad (1)$$

where  $(\bar{x}, \bar{y})$  are the center of mass of  $s$ . Denote  $\lambda_1$  and  $\lambda_2$  as the two eigenvalues of the normalized inertia matrix of  $s$ , with  $\lambda_1 \geq \lambda_2$ ,  $a$  as its area,  $p$  as its perimeter,  $u(s)$  to be the gray value of the pixels only contained by  $s$ ,  $\mu_s$  and  $\sigma_s$  as the mean and the standard deviation of gray values of all

<sup>1</sup>The codes of FLST are included in the free software MegaWave, and can be downloaded at <http://megawave.cmla.ens-cachan.fr/>.

pixels inside  $s$ ,  $s^r, r \in [1, \dots, M]$  to be the  $r$ -order ancestor of shape  $s$  in the tree, and  $a_{max}, a_{min}$  to be two thresholds on shape area, the attributes used for characterizing shape  $s$  are given in Table I. As a summarization, the shape attributes of  $s$  can be written as,

$$f(s) = [\epsilon, \kappa_e, \kappa_c, \gamma, \alpha, \theta], \quad (2)$$

each component of which is invariant or robust to different geometrical variations remarked in Table I, respectively.

TABLE I. ATTRIBUTES FOR CHARACTERIZING A SHAPE  $s$ . FOR THE NOTATION OF SYMBOLS, PLEASE REFER TO THE TEXT FOR DETAILS.

attribute	computation	invariance
Elongation	$\epsilon = \lambda_2 / \lambda_1$	similarity
Ellipse-compactness	$\kappa_e = a / (4\pi\sqrt{\lambda_1\lambda_2})$	affine
Circle-compactness	$\kappa_c = 4\pi p^2 / a$	similarity
Contrast	$\gamma = (u(s) - \mu_s) / \sigma_s$	affine
Scale ratio	$\alpha = M a / (\sum_{r=1}^M a(s^r))$	affine
Normalized area	$\theta = (\ln a - \ln a_{min}) / (\ln a_{max} - \ln a_{min})$	affine

SITA uses the marginal distributions of these attributes of all the shapes contained in an image as features for texture analysis.

## B. Shapes co-occurrence patterns (SCOPs)

As explained in previous section, shapes relationship on the hierarchical structure of the topographic map plays an important role in texture analysis. In spatial domain, such shapes' relationship corresponds to local co-occurrence structures in an image. In the topographic map, a hierarchical tree, the main relationships are sibling and conclusion, corresponding to small branches on the tree. If we take such small branches as textons, texture modeling is then to investigate the branches arrangement laws on the tree.

In our context, we define *shape co-occurrence patterns* as *simple and local common branches on the tree of shapes*, which reflects specific spatial organization in the topographic map. Observe that, though the shape co-occurrence patterns are local on the tree of shapes, they may correspond to large spatial areas of pixels in the images, which enables us to consider highly geometrical and complex texture elements, such as co-occurred elongated or sharp edges.

1) *Common SCOPs*: Denoting  $s^r$  as the  $r$ -order ancestor of shape  $s$  on the tree,  $s'$  as the sibling of  $s$  and  $s^\tau$  to be the grand-ancestor of  $s$  with  $\tau > r$ , in this paper, we use following four common SCOPs as

$$\begin{aligned} \text{single shape } (P_1) &: s, \\ \text{shape-ancestor } (P_2) &: s \rightarrow s^r, \\ \text{shape-ancestor-grandancestor } (P_3) &: s \rightarrow s^r \rightarrow s^\tau, \\ \text{shape-ancestor-sibling } (P_4) &: s \rightarrow s^r \leftarrow s', \end{aligned}$$

respectively.

Fig.3 illustrates two common SCOPs  $P_3$  and  $P_4$ , which are very popular and efficient in texture presentation as they encode rich spatial co-occurrence information.

Two parameters for constructing the SCOPs are the interval order  $r$  between the shape and its ancestor, and the levels of the interval  $\frac{\tau}{r}$ .

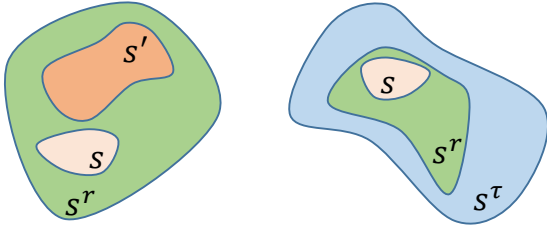


Fig. 3. The shape-ancestor-sibling ( $P_4 : s \rightarrow s^r \leftarrow s'$ ) (a) and shape-ancestor-grandancestor  $P_3 : s \rightarrow s^r \rightarrow s^\tau$  (b).

a) *The interval  $r$* : The interval  $r$  controls the correlation between two adjacent shapes in the tree. Observe that a too small  $r$  makes no sense as the two adjacent shapes will be almost the same, while a too big  $r$  will not be local enough and results in weak descriptive ability. In our context, we proposed to use  $r$ , that the difference between the areas of the pair of shapes are larger than the perimeter  $p$  of shape  $s$ ,

$$r = \min\{i; a_{s^i} - a_s > p(s)\}, \quad (3)$$

where  $s^i$  denotes the  $i$ -th order ancestor of the shape  $s$ .

Note that  $r$  in fact approximates the number of blur around objects in nature images. The final choice of the interval  $r$  is the average value given by Eq. (3) on all shapes in the images, which corresponds to the averaging blur in textures.

b) *The number of cascaded ancestors  $\tau$* :  $\tau$  has similar effect for texture description as  $r$ . Considering the casual relationship between two cascaded shapes in the tree structure,  $\tau$  is set to be integer times of  $r$ . In our experiments,  $\tau$  is set to be  $2r$  to achieve better performances.

2) *Learn SCOPs with K-means*: As a local common branch on the tree, each SCOP  $P_k$  actually contains tremendous image or shape realizations described by branch attributes. For the common SCOPs ( $P_k$ ) $_{k=1, \dots, 4}$ , the attributes are

$$\begin{aligned} \text{attributes of } P_1 &: f(s), \\ \text{attributes of } P_2 &: [f(s), f(s^r)] \\ \text{attributes of } P_3 &: [f(s), f(s^r), f(s^\tau)], \\ \text{attributes of } P_4 &: [f(s), f(s^r), f(s')], \end{aligned}$$

where  $f(s)$  is the attribute associated to  $s$ .

In order to simplify the problem, we need to quantize the shape realizations of each SCOP into limited number of clusters, named *words* of a SCOP. Such words can be learned from the tree of shapes of a given set of texture images, as illustrated in Figure 1. In this paper, we use K-means clustering algorithm to achieve this goal. Figure 4 shows several learned words for each SCOP.

### C. Texture Descriptor from SCOPs

So far, we have the learned dictionary of visual words for several common SCOPs, in what follows, we then develop texture descriptor from them.

Following the bag-of-words strategy, each texture can be modeled by encoding the ensemble of branches of shapes it

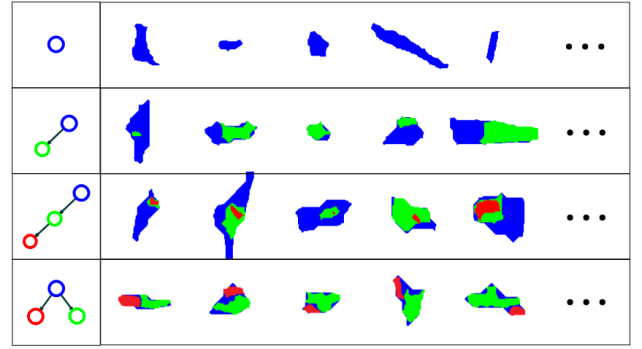


Fig. 4. 4 common SCOPs with several learned *words* in corresponding row.

contained in the images according to the learned dictionary of SCOPs. Thus the model of a texture  $t$  is,

$$t \sim [h^t(P_1), h^t(P_2), \dots, h^t(P_K)],$$

where  $h^t(P_k)$  is the histogram of the SCOP  $P_k$  of texture image  $t$ , and  $K = 4$  in our case.

As a summary, the whole pipeline of SCOP-based texture analysis method is illustrated in Figure 1.

## IV. EXPERIMENTAL RESULTS

This section first explains the experimental setting and how to compare texture images using the model introduced in the previous section. We then investigate the performances of the resulting comparison scheme by confronting it with state-of-the-art texture descriptors [3], [4], [8], [14], on two challenging texture data sets for invariant texture recognition. We replicate, as closely as possible, the experiments described by [3], [4], [8], [14]. These consist of two tasks, retrieval and classification, applied to two data sets UIUC [3] and UMD [21].

### A. Experimental setup

1) *Comparing textures*: Given two texture samples  $u$  and  $v$ , associated with models  $h^u = [h^u(P_1), h^u(P_2), \dots, h^u(P_K)]$  and  $h^v = [h^v(P_1), h^v(P_2), \dots, h^v(P_K)]$  respectively, we choose to compare histograms through the *histogram intersection kernel* (HIK), which is widely used for comparing bag-of-words model,

$$Dist(h^u, h^v) = \sum_{k=1}^K \sum_{i=1}^{n_k} \min(h^u(P_k)[i], h^v(P_k)[i]). \quad (4)$$

where  $h^u(P_k)[i]$  is the  $i$ -th bin of the histogram  $h^u(P_k)$  and  $n_k$  is the number of bins.

Moreover, we also use the approximated geodesic distance to compare two texture samples, as it has demonstrated its capability to improve recognition performances [4]. More precisely, based on the computed pairwise distance matrix using HIK as in Eq. (4), we construct the  $k$ -nearest neighbor graph of the samples (two vertices are connected if one is among the  $k$ -nearest neighbors of the other). The geodesic distance between two points is then approximated by the shortest path between them on the graph. This shortest path may for instance be computed using Dijkstra's algorithm.

2) *Data sets*: We use two dataset, the UIUC [3] database and the UMD database [21]. Both of them contain 25 texture classes, each one being composed of 40 samples (*i.e.* 1000 samples altogether). The UIUC database is of size  $640 \times 480$ , and the samples are subject to drastic viewpoint changes, contrast changes or even non-rigid deformations inside each class. See Figure 6 for an example. In contrast, the UMD Database shows strong viewpoint and scale changes, and significant contrast differences. A significant proportion of this database is made of textures consisting in the repetition of objects. Moreover, the resolution of the images in UMD Database is  $1280 \times 960$ .

3) *Parameters setting*: In our experiments, the parameter  $r$  is estimated and set to be 5 and  $\tau = 2r$  for both the two texture data sets. The minimal area  $a_{min}$  and maximal area  $a_{max}$  of the shapes are chosen as 3 and 8000 respectively. For the 4 common SCOPs, the numbers of clusters in K-means algorithm are set to be  $\{100, 200, 300, 300\}$  in the bag-of-words strategy, which are used to construct the histogram representation of each texture sample.

### B. Retrieval experiments

The retrieval experiment consists in using one sample of the database as a query to retrieve the  $N_r$  most similar samples in the data set. For evaluation, the average number of correctly retrieved samples (generally called recall) when the query spans the whole database is drawn as a function of  $N_r$ .

Figures 5 (a) and (b) show the retrieval results of the proposed SCOPs-based descriptors with geodesic (SCOP+Geo) on the UIUC and UMD database. Observe that the performances of SCOPs+Geo are largely better than those of [4], [8] on both data sets. More precisely, when the recall number is 39 the retrieval rates of the proposed method are respectively 89.4% and 89.3% on the UIUC and UMD database, comparing with 61.2% and 66.1% of the scattering transform features with geodesics (ST+GEO) and 78.5% and 87.0% of the SIAT approach with geodesics (SITA+Geo).

It is interesting to note that the proposed SCOP+Geo method achieve extremely high retrieval rates on textures with 3-D deformations, for instance 100% it can retrieve the Plaid class in UIUC data set with 100% precision, which is reported to be very difficult to handle, see Figure 6 for an example<sup>2</sup>.

### C. Classification experiments

For the classification experiment, the SVM classifiers with the histogram intersection kernel (HIK) is be used to classify the images. Different number of samples (5, 10 and 20 in our experiments) are extracted from each class in the dataset and used as a training set for classification. The final classification rate are the average of that achieved by 200 times random splitting of the data sets.

On the UIUC data set, when the training size is 20, the scattering transform with “log, scale average, multi-scale train” strategy outperforms other methods, while the proposed SCOP gives comparable results. However, when the training size decreases (e.g. from 20 to 5), our SCOP-based method

performs much better than all other existing approaches (about 4%). Similar performances are obtained on UMD data set. A quantitative comparison on classification performances of several recent texture analysis methods are given in Table II.

TABLE II. COMPARISON OF CLASSIFICATION ACCURACY ON UIUC [3] DATABASE AND UMD DATABASE [14].

Training size	UIUC			UMD		
	5	10	20	5	10	20
Lazebnik [3]	87.2	92.6	96.0	84.7	89.1	92.5
WMFS [14]	93.4	97.0	98.6	93.4	97	98.7
SITA [4]	91.5	95.0	97.5	95.1	98.7	99.1
ST [8]	93.3	97.8	99.4	96.3	98.9	99.7
SCOP	95.8	97.9	98.9	97.1	98.7	99.4
SCOP+Geo	97.7	98.4	98.7	97.9	98.8	99.3

So far, we can conclude that the proposed texture analysis approach with SCOPs can achieve comparable classification performance with the-state-of-the-art texture analysis methods when using a large number of training samples, while, it largely outperforms all existing texture analysis approaches when a small set of training samples are available, especially for image retrieval, where only one training sample is at hand.

## V. CONCLUSION

This paper introduced a texture analysis framework by investigating the co-occurrence patterns of shapes via bag-of-words strategy. The proposed method inherits the good properties from the shape-based texture method, and provides a more flexible way to consider shape relationships and high-order statics on the tree. The experiments of texture retrieval and classification on two challenging texture databases demonstrate that the proposed method has strong ability to depict geometrical aspects of textures and is very robustness to the changes of viewpoints and illuminations. Remark that the SITA [4] is actually an special case of the proposed method, where the shape attributes are supposed to be independent and only their marginal distributions are adopted for texture characterization. As the method shows great potentials in depicting the geometrical aspects of textures as well as being robust to scale and illumination changes, it may be a good tool for analyzing high-resolution remote sensing images, where thousands of objects with various sizes and colors emerge in the same scene [22]. Another possible work in the future is to use the proposed texture features as visual cues for achieving image segmentation.

## ACKNOWLEDGMENT

The authors would like to thank Stephane Mallat, who kindly shares us with his scattering transform texture analysis codes. This work was partially founded by NSFC project No.91338113.

## REFERENCES

- [1] T. K. Leung and J. Malik, “Representing and recognizing the visual appearance of materials using three-dimensional textures,” *Int. J. Comput. Vision*, vol. 43, no. 1, pp. 29–44, 2001.
- [2] J. Portilla and E. P. Simoncelli, “A parametric texture model based on joint statistics of complex wavelet coefficients,” *Int. J. Comput. Vision*, vol. 40, no. 1, pp. 49–70, 2000.
- [3] S. Lazebnik, C. Schmid, and J. Ponce, “A sparse texture representation using local affine regions,” *IEEE Trans. PAMI*, vol. 27, no. 8, pp. 1265–1278, 2005.

<sup>2</sup>More results can be founded in the supplemental materials.

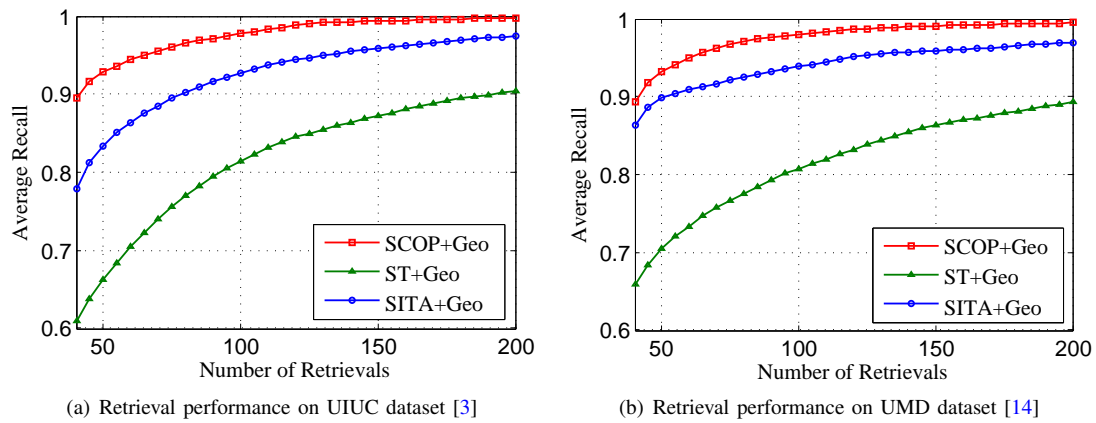


Fig. 5. Average retrieval performances of different methods on UIUC data set (left) and UMD data set (right). In both figures, the red curves indicate our retrieval results, while the blue and green ones illustrate the results of SITA [4] and Scatter Transform (ST) [8].

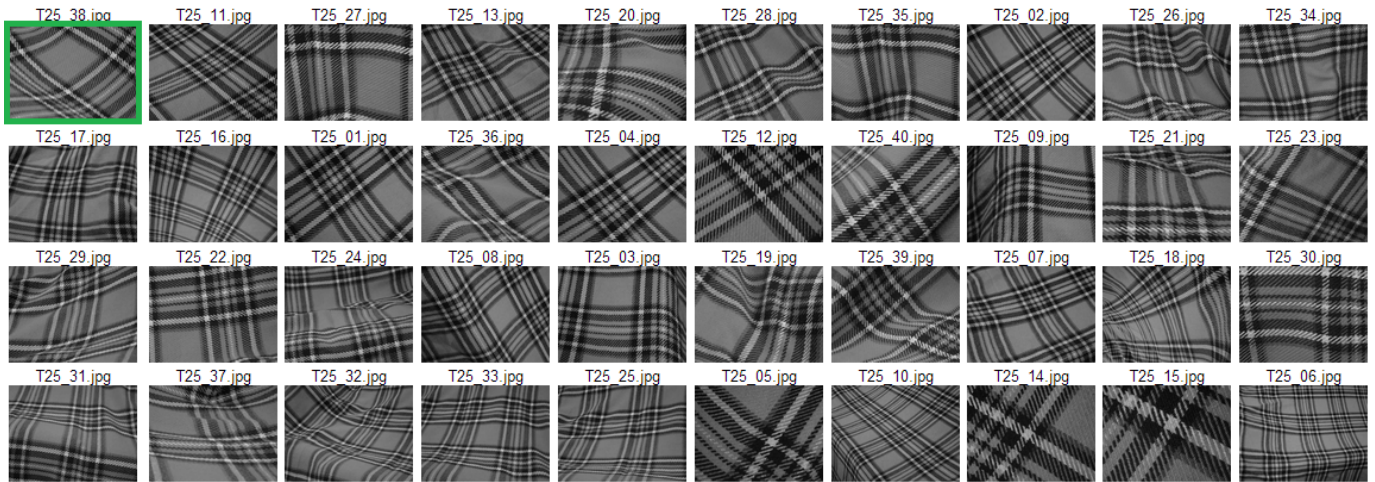


Fig. 6. A retrieval result on the Plaid class in UIUC data set using the proposed method. The image in green rectangle is the query image and the rests are the retrieved ones ranked by their similarities.

- [4] G.-S. Xia, J. Delon, and Y. Gousseau, "Shape-based invariant texture indexing," *Int. J. Comput. Vision*, vol. 88, no. 3, pp. 382–403, 2010.
- [5] P. Kim, Steenstrup, S.-S. Kristoffer, Z. Andrew, and I. Christian, "Shape index descriptors applied to texture-based galaxy analysis," in *Proc. ICCV*, 2013.
- [6] R. Li and E. Adelson, "Example-based facade texture synthesis," in *Proc. CVPR*, 2013.
- [7] T. Matthews, M. Nixon, and M. Niranjan, "Enriching texture analysis with semantic data," in *Proc. CVPR*, 2013.
- [8] L. Sifre, S. Mallat, and E. N. S. DI, "Rotation, scaling and deformation invariant scattering for texture discrimination," in *Proc. CVPR*, 2013.
- [9] S.-C. Zhu, C.-E. Guo, Y. Wang, and Z. Xu, "What are textons?," *Int. J. Comput. Vision*, vol. 62, no. 1-2, pp. 121–143, 2005.
- [10] G. Peyré, "Texture synthesis with grouplets," *IEEE Trans. PAMI*, vol. 32, no. 4, pp. 733–746, 2009.
- [11] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images," *Int. J. Comput. Vision*, vol. 62, no. 1-2, pp. 61–81, 2005.
- [12] P. Monasse and F. Guichard, "Fast computation of a contrast-invariant image representation," *IEEE Trans. Image Processing*, vol. 9, no. 5, pp. 860–872, 2000.
- [13] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. PAMI*, vol. 24, no. 7, pp. 971–987, 2002.
- [14] Y. Xu, X. Yang, H. Ling, and H. Ji, "A new texture descriptor using multifractal analysis in multi-orientation wavelet pyramid," in *Proc. CVPR*, pp. 161–168, IEEE, 2010.
- [15] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Systems, Man and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.
- [16] L. Davis, "Polarograms: a new tool for image texture analysis," *Pattern Recognition*, vol. 13, no. 3, pp. 219–223, 1981.
- [17] M. Pietikainen, T. Ojala, and Z. Xu, "Rotation-invariant texture classification using feature distributions," *Pattern Recognition*, vol. 33, no. 1, pp. 43–52, 2000.
- [18] B. Julesz, "Textons, the elements of texture perception, and their interactions," *Nature*, vol. 290, no. 5802, pp. 91–97, 1981.
- [19] F. Lafarge, G. Gimel'farb, and X. Descombes, "Geometric feature extraction by a multimarked point process," *IEEE Trans. on PAMI*, vol. 32, no. 9, pp. 1597–1609, 2010.
- [20] V. Caselles, B. Coll, and J.-M. Morel, "Topographic maps and local contrast changes in natural images," *Int. J. Comput. Vision*, vol. 33, no. 1, pp. 5–27, 1999.
- [21] Y. Xu, H. Ji, and C. Fermuller, "Viewpoint invariant texture description using fractal analysis," *Int. J. Comput. Vision*, vol. 83, no. 1, pp. 85–100, 2009.
- [22] G.-S. Xia, W. Yang, J. Delon, and Y. Gousseau, "Structural high-resolution satellite image indexing," in *ISPRS TC VII Symposium-100 Years ISPRS*, vol. 38, pp. 298–303, 2010.