Texture Classification using a Linear Configuration Model based Descriptor

Yimo Guo
guoyimo@ee.oulu.fi

Guoying Zhao
gyzhao@ee.oulu.fi

Matti Pietikäinen
mkp@ee.oulu.fi

Machine Vision Group
University of Oulu
OULU, Finland

Abstract

We investigate rotation invariant image description and develop a linear model based descriptor namely MiC, which is suited to modeling microscopic configuration of images. To explore multi-channel discriminative information of both the microscopic configuration and local structures, the feature extraction process is formulated as an unsupervised framework that consists of: 1) the configuration model to encode image microscopic configuration; and 2) local patterns to describe local structural information. In this way, images are represented by a novel feature: local configuration pattern (LCP). We evaluate the performance of the proposed method by classifying textures present in three challenging texture databases: Outex_TC_00012, KTH-TIPS2 and Columbia-Utrecht (CUReT). The encouraging results show that LCPs is highly discriminative.

1 Introduction

Image textures are inherent and complex visual patterns that reflect the information of gray level statistics, spatial distribution, synthetic structure and so on. Texture analysis aims to interpret and understand these real-world visual patterns, which can be broadly used in image filtering, classification, segmentation, indexing and synthesis. As one of the fundamental topics, texture classification has been applied to many practical vision systems, such as the classification of regions in satellite images and detection of defects in industrial surface inspection. It has also been used in biomedical image analysis to diagnose tissues for possible diseases, for example, cancer or tumours. Because of the wide variety of potential applications, it receives lots of research interest from both academia and industry.

The perception of texture plays an important role in the human visual system and is considered to be a relatively easy task for human beings. However, it becomes challenging when the analysis is made by computational algorithms. Basically, texture classification can be concluded as the problem of classifying images according to textural cues, that is, categorizing a texture image obtained under certain illumination and viewpoint condition as belonging to one of the pre-learned texture classes. Therefore, it would mainly pass through two steps: image representation or description and classification. In this paper, we focus
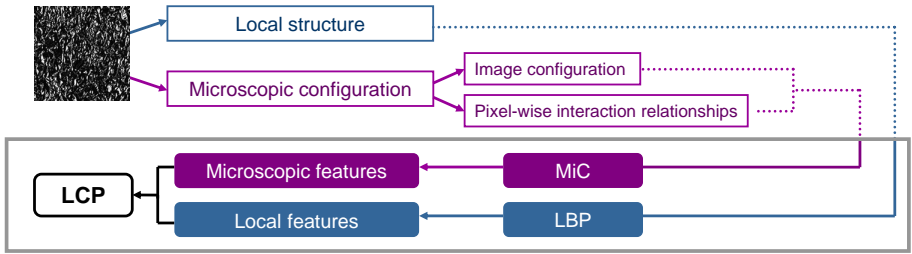
Figure 1: The feature extraction framework of LCPs.

on the feature extraction part that aims to extract effective patterns to distinguish different textures.

2 Related work

Texture classification has been extensively studied in the past decades and numerous feature extraction approaches have been proposed. Representative methods include scale-invariant feature transform (SIFT) and related methods [13, 14, 27], local binary patterns (LBPs) and its extensions [1, 2, 8, 9, 12, 16, 18], grey level difference or co-occurrence statistics [26], and methods based on multi-channel filtering or wavelet decomposition [15, 20, 22]. Among these methods, local features have performed well in real-world applications, such as LBP, SIFT and Histogram of Oriented Gradients (HOG) [5]. These descriptors explore local structures, which contain low-level characterization (distribution of color, coarseness, contrast, etc.) and basic visual elements (dots, lines, circles, etc.). In particular, LBP [18], an efficient local feature, was reported to detect local structures, such as lighting spots and edges. This kind of local structures play an important role in human texture perception. However, LBP based local features might be statistically unreliable due to the limitations brought by the encoding scheme that quantizes local information by estimating intensity properties, such as the average and variance, of neighboring pixels. On the other hand, image configuration has been also investigated for discrimination. For example, R. Chellappa *et al.* [4] presented the Gaussian Markov random field (GMRF) to model texture image patterns. This framework was later developed for rotation invariant texture classification [7, 11]. To learn representative structural configuration from texture images, Varma *et al.* proposed texton methods based on the filter response space and local image patch space [23, 24, 25]. Their performance shows that properties of texture, other than the statistical information of local binary patterns, contain powerful discriminative information as well.

In this paper, we propose the descriptor MiC that encodes image microscopic configuration by a linear configuration model. The final local configuration pattern (LCP) feature integrates both the microscopic features represented by optimal model parameters and local features represented by pattern occurrences, as shown in Figure 1. To be specific, microscopic features capture image microscopic configuration which embodies image configuration and pixel-wise interaction relationships by a linear model. The optimal model parameters are estimated by an efficient least squares estimator. To achieve rotation invariance, which is a

desired property for texture features, Fourier transform is applied to the estimated parameter vectors. Finally, the transformed vectors are concatenated with local pattern occurrences to construct LCPs. As this framework is unsupervised, it could avoid the generalization problem suffered by other statistical learning methods.

3 Local configuration pattern (LCP)

LCP feature decomposes the information architecture of images into two levels, as shown in Figure 1: 1) local structural information; 2) microscopic configuration information that involves image configuration and pixel-wise interaction relationships. For local structural information, we utilize the LBP in feature extraction framework, while a microscopic configuration model is developed to explore microscopic configuration information.

3.1 LBP-based local features

LBPs are usually extracted in a circularly symmetric neighborhood by comparing each image pixel with its neighborhood, expressed by:

$$LBP(P,R) = \sum_{i=0}^{P-1} u(g_i - g_c)2^i,$$ (1)

where P is the number of neighboring samples and R is the radius of neighborhood. g_i denotes the intensity value of neighboring pixel i ($i = 0,...,P-1$), g_c is the intensity value of the center pixel, and $u(x)$ is a step function with $u(x) = 1$ if $x \geq 0$ and $u(x) = 0$ otherwise. The intensities of neighboring pixels which do not fall exactly on the image grid are obtained by bi-linear interpolation. The coordinate of neighboring sample i can be determined by $(x_i, y_i) = (x_c + R\cos(2\pi i/P), y_c - R\sin(2\pi i/P))$, where (x_c, y_c) denotes the coordinate of the center pixel. Those patterns that have at most two bit-wise transitions in circular binary representation are defined to be 'uniform' [18]. The rotation invariant patterns can be obtained by rotating each bit pattern circularly to the minimum value. Therefore, a rotation invariant uniform pattern can be defined as:

$$LBP^{riu2}(P,R) = \begin{cases} \sum_{i=0}^{P-1} u(g_i - g_c), & \text{if } U(\text{LBP(P,R)}) \leq 2 \\ P+1, & \text{otherwise} \end{cases},$$ (2)

where U is the uniformity measure to define uniform patterns, decided by the number of spatial transitions.

Although image representation by LBP-based methods could increase the robustness against illumination variation, the capability of encoding image configuration and pixel-wise relationships might be weakened as they quantize gray-level differences to two binary levels. For example, with the LBP representation [18], the patterns shown in Figure 2 (a) and (b) would be classified into the same class. But the textural surfaces they represent are quite different from each other, which means they probably belong to different classes. More specifically, the texture class, with the pattern in Figure 2 (a), probably has smoother surfaces as intensity variations between the center pixel and neighborhood are small; while the texture class, with the pattern in Figure 2 (b), should have a rough surface containing sharp edges due to higher intensity variations between the center pixel and neighborhood. Therefore, the classification might be inaccurate due to the lack of magnitude information in features. To
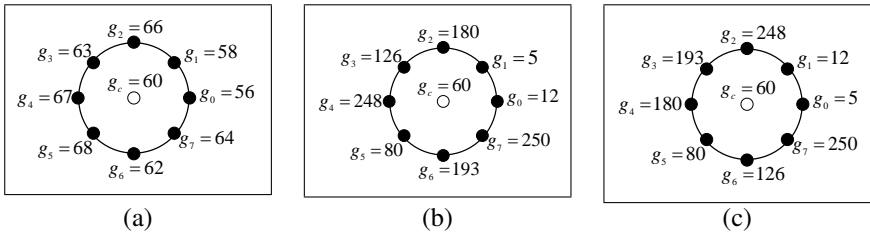
Figure 2: Patterns in (a) and (b) would be considered as the same pattern type by LBP, though corresponding textural surfaces might be quite different from each other. By incorporating LBP with local variance information, patterns in (a) and (b) could be distinguished, while patterns in (b) and (c) would still be considered as the same pattern type because of the same variance. But they are different in configuration, which is not due to the rotation but underlying textural properties. In these figures, g_c denotes the intensity of the center pixel and g_i (i=0,...,7) denotes the intensities of neighborhood.

solve this problem, local variance information (VAR), which can be calculated by Equation 3, was adopted together with the LBPs [18] to form the joint histogram: LBP^{riu2}/VAR.

$$VAR = \frac{1}{P}\sum_{i=0}^{P-1}(g_i - \mu)^2, \qquad (3)$$

where μ is the average intensity value of all neighboring pixels. However, the joint histogram of LBP and VAR can not fully solve the problem. The classification might be misled without considering the relationships among neighborhood intersection. Taking the patterns in Figure 2 (b) and (c) for example, which would be considered as the same pattern type according to Equations 2 and 3, they are actually two patterns with different textural properties. Therefore, in the proposed feature extraction framework, we investigate both the image configuration and pixel-wise interaction relationships as discrimination for effective image representation.

3.2 Microscopic configuration modeling (MiC)

In this subsection, we concentrate on the MiC, a microscopic configuration model developed to obtain microscopic features which reflect textural properties.

Modeling of microscopic configuration. To model the image configuration with respect to each pattern, we estimate optimal weights, associating with intensities of neighboring pixels, to linearly reconstruct the central pixel intensity. This can be expressed by:

$$E(a_0,...,a_{P-1}) = |g_c - \sum_{i=0}^{P-1} a_i g_i|. \qquad (4)$$

In this formula, g_c and g_i denote intensity values of the center pixel and neighboring pixels of a particular pattern type respectively, a_i $(i = 0,...,P-1)$ are weighting parameters associated with g_i, and $E(a_0,...,a_{P-1})$ is the reconstruction error regarding model parameters a_i $(i=0,...,P-1)$. To minimize the reconstruction error for each pattern, optimal parameters are determined by the least squares estimation [21].

Least squares estimation for optimal model parameters. Given an image I, suppose the occurrence of a particular pattern type L is N_L, which means there are N_L pixels in I

with the pattern type L. The label code can be calculated by Equation 1 or 2, depending on what kind of patterns are of interest. We denote the intensities of these N_L pixels as $c_{L,i}$ $(i = 0, ..., N_L - 1)$, and organize them into a vector:

$$C_L = \begin{pmatrix} c_{L,0} \\ c_{L,1} \\ \cdot \\ \cdot \\ \cdot \\ c_{L,N_L-1} \end{pmatrix}. \tag{5}$$

The intensities of their neighboring pixels $v_{i;0}, ..., v_{i;P-1}$ $(i = 0, ..., N_L - 1)$ can thus be organized as:

$$V_L = \begin{pmatrix} v_{0;0} & v_{0;1} & \cdots & v_{0;P-1} \\ v_{1;0} & v_{1;1} & \cdots & v_{1;P-1} \\ & & \cdot & \\ & & \cdot & \\ & & \cdot & \\ v_{N_L-1;0} & v_{N_L-1;1} & \cdots & v_{N_L-1;P-1} \end{pmatrix}. \tag{6}$$

In order to minimize the reconstruction error in Equation 4, the unknown parameters a_i $(i = 0, ..., P - 1)$ are constructed as a column vector:

$$A_L = \begin{pmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_{P-1} \end{pmatrix}. \tag{7}$$

In this way, the problem to be solved becomes a least-squares problem $C_L = V_L A_L$. When the system is over-determined, optimal parameter vector A_L is determined by:

$$A_L = (V_L^T V_L)^{-1} V_L^T C_L. \tag{8}$$

Otherwise, when $N_L \leq P$, the pattern L rarely occurs, so it would be considered as a non-reliable pattern to serve as a feature. In this case, each entry of the parameter vector will be set to zero. By considering the pixel-wise interaction, the parameter vectors of different patterns shown in Figure 2 (b) and (c) would not be the same since each entry of A_L changes according to the relative neighborhood positions. Therefore, the obtained feature is capable of capturing both the information of magnitude and pixel-wise interactions, which makes it more discriminative than the previous representation.

In the area of texture analysis, rotation invariant analysis is a widely studied problem, which aims at providing with texture features that are invariant to rotation angle of the input image. To produce rotation invariant features, we apply 1D Fourier transform to the estimated parameter vector A_L. The transformed vector can be expressed by:

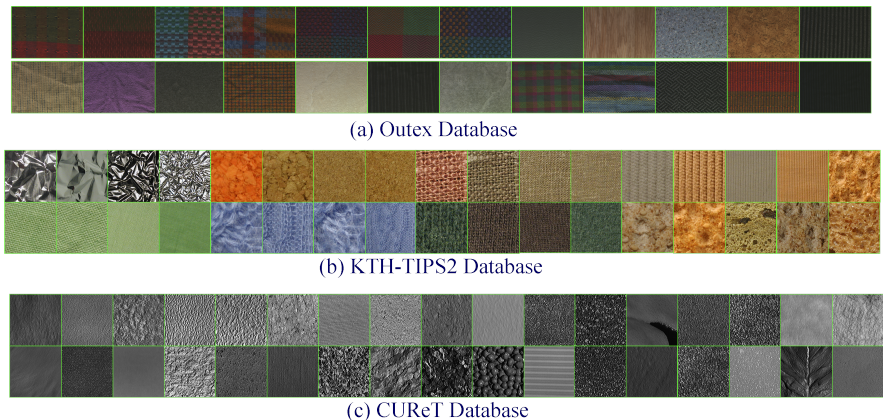$$H_L(k) = \sum_{i=0}^{P-1} A_L(i) \cdot e^{-j2\pi ki/P}, \tag{9}$$

(a) Outex Database

(b) KTH-TIPS2 Database

(c) CUReT Database

Figure 3: Samples of Outex, KTH-TIPS2 and CUReT databases.

where $H_L(k)$ is the kth element of H_L and $A_L(i)$ is the ith element of A_L. Although image rotation would lead to cyclic translations of A_L, Fourier transform is invariant to this kind of translations so that H_L could achieve rotation invariant property. The magnitude part of vector H_L is taken as the MiC feature, which is defined by:

$$|H_L| = [|H_L(0)|; |H_L(1)|; ...; |H_L(P-1)|]. \tag{10}$$

Considering that $|H_L|$ encodes the image configuration and pixel-wise interaction relationship of each specific pattern, it together with pattern occurrences of local binary patterns would construct a complementary feature for both the discrimination of microscopic configuration and local structures. In this way, the final feature is:

$$LCP = [[|H_0|; O_0]; [|H_1|; O_1]; ...; [|H_{q-1}|; O_{q-1}]], \tag{11}$$

where $|H_i|$ is calculated by Equation 10 with respect to the ith pattern of interest, O_i is the occurrence of the ith local pattern of interest (i.e., the LBP), and q is the total number of patterns of interest. Moreover, multi-scale analysis can be achieved by combining LCPs with different radii and neighboring samples.

4   Experiments

We evaluate the performance of the proposed method on three comprehensive databases by classifying textures in different scenarios: Outex database [17] for rotation invariant texture classification, KTH-TIPS2 database [3] for material categorization and Columbia-Utrecht (CUReT) database [5] for classification under different views and illuminations, as seen in Figure 3. LCP is also compared with several state-of-the-art texture classification approaches on all these databases.

4.1   Rotation invariant texture classification

We conduct experiments on the challenging test suite Outex_TC_00012. It consists of 9,120 images representing 24 different textures imaged under different illumination conditions and

rotation orientations [17]. The test set contains 20 training images for each texture class. The training images are taken under single orientation whereas different orientations are present in 8,640 testing images. The classification rates over all test images are listed in Table 1, which are derived from the setup by using the nearest neighbor classifier. For all the compared methods, multi-scale analysis is implemented. The number of neighboring samples for LCP is fixed to eight, while the number of neighboring samples for other approaches increase as the radius enlarges [2]. LBP^{riu2} is introduced as patterns of interest for LCP. We adopt the nearest neighbor classifier for DLBP without any preprocessing. In this paper, chi-square is used as the similarity measure for LBP, while L1 is used as the distance for LBP-HF, DLBP and LCP since these feature vectors yield better performances using L1 distance.

From the experimental results, it is observed that LCPs, which could provide the highest classification accuracy among all the compared methods, is highly discriminative. The classification accuracy of the proposed method can also be further improved with the multi-scale setting $R_2 + R_3 + R_4$ (i.e., up to 0.943), as shown in Table 1. Although the classification performance can be improved by multi-scale analysis for all approaches, LCP significantly outperforms others even with single scales, for example, in the case of $R = 3$. For further comparison, we refer to texton based methods [24, 25] and the adaptive LBP (ALBP) [10]. The texton method got 0.917 by Joint method and 0.927 by MR8 on this test suite, while ALBP got 0.894 when $R = 3$ as the best result in single scale cases.

Methods	LBP^{u2} [18]	LBP^{riu2} [18]	LBP-HF [2]	DLBP	LCP
R_1	0.566	0.646	0.773	0.560	0.684
R_2	0.578	0.791	0.873	0.687	0.881
R_3	0.450	0.833	0.896	0.754	0.923
$R_1 + R_2$	0.595	0.821	0.894	0.778	0.840
$R_1 + R_3$	0.512	0.883	0.917	0.820	0.880
$R_2 + R_3$	0.513	0.857	0.915	0.837	**0.927**
$R_1 + R_2 + R_3$	0.539	0.870	0.925	0.849	0.903
$R_2 + R_3 + R_4$					**0.943**

Table 1: Classification rates of different approaches on the Outex database. R_i means radius value $R = i$.

4.2 Material categorization

The proposed method is also assessed on the KTH-TIPS2 database [3]. This database contains 11 different classes of texture materials, each class has 4 different samples, and each sample was imaged at 9 different scales and 12 different illumination and pose conditions. In this experiment, one sample is randomly selected from each class to serve as the training set, other images are used as the testing set. This is repeated 500 times with different combinations of training and testing sets. The setting for different approaches follows the experiment conducted on the Outex database, except that uniform patterns are used as patterns of interest considering that rotation invariance does not benefit much on this database.

The average classification rates of different approaches are listed in Table 2. From experimental results, LCP has obvious superiority among the compared methods in all cases. The capability of LCP and its robustness against variant illumination and pose conditions is demonstrated. It can be found that classification accuracies obtained by using uniform patterns are consistently higher than those of using rotation invariant uniform patterns, which matched the observations in [2]. In addition, the performance of LBP^{riu2} is consistently

Methods	LBP^{u2} [?]	LBP^{riu2} [?]	LBP-HF [?]	DLBP	LCP
R_1	0.528	0.482	0.524	0.458	0.592
R_2	0.511	0.494	0.535	0.460	0.579
R_3	0.502	0.481	0.516	0.459	0.570
$R_1 + R_2$	0.536	0.502	0.539	0.456	0.592
$R_1 + R_3$	0.542	0.507	0.545	0.468	0.593
$R_2 + R_3$	0.514	0.508	0.542	0.458	0.574
$R_1 + R_2 + R_3$	0.536	0.514	0.548	0.461	**0.608**
MR8[24]	0.455				

Table 2: Classification rates of different approaches on the KTH-TIPS2 database.

lower probably because different orientations have been present in the training data.

4.3 Texture classification under variant imaging conditions

To evaluate the performance under variant imaging conditions, we perform experiments on the CUReT database. This database includes 61 different real-world textures, and each class has 205 images obtained at different viewpoints and illumination angles [6]. To compare the results, two experiments are performed on this database to deal with different classification problems considering both 40 and all 61 textures, as carried out in [19, 23]. Half of the 92 images for each texture are used for training and the other half for testing. For example, it contains 2,806 images respectively taken under varying viewpoint and illumination with a viewing angle less than 60 degrees for 61 textures. The training (and testing) set are selected randomly and an average accuracy is computed by 500 trial runs. These images are cropped to resolution 200×200 [1].

For comparison, the results obtained by LBP^{riu2} with contrast VAR are also provided. Three scales R_1, R_3, R_5 with neighboring samples 8, 16 and 24 are adopted for LBP, LBP^{riu2}, LBP/VAR and LBP-HF. DLBP and LCP follow the setting of R_1, R_3 and R_5 with eight neighboring samples. The average classification rates of different approaches are listed in Table 3. It is observed that LBP^{riu2} gives even better classification rates than LBP^{riu2}/VAR which was proposed as a joint distribution of LBP and VAR to represent local spatial information more completely. As contrast basically encodes the variance of local neighborhood, it would be sensitive to variant conditions such as the illumination. When texture images contain clear orientation, as in this database, the isotropic measurement may fail to provide effective image representation [10]. However, the proposed method could still achieve the highest classification accuracies among all compared approaches. The obvious superiority of LCP in these two experiments could strongly verify its robustness and capability under variant imaging conditions.

5 Conclusions

We propose a statistical descriptor MiC for texture image description, which is based on a linear configuration model, and combine it with the LBP to achieve LCP descriptor. The derived LCPs, which considers both the microscopic configuration and local structure information, consist of two parts: 1) MiC to encode linear relationships among neighboring

---

[1]Available at: http://www.robots.ox.ac.uk/~vgg/research/texclass/index.html

	LBP [19]	LBP^{riu2} [19]	LBP^{riu2}/VAR	LBP-HF	DLBP	MR8 [19]	LCP
40	0.916	0.973	0.940	0.964	0.934	0.963	**0.977**
61	0.900	0.966	0.957	0.968	0.959	0.961	**0.979**

Table 3: Classification rates on the CUReT database with different texture classes: 40 and 61.

pixels; 2) local descriptor to explore local shape information and construct pattern occurrence histograms. The optimal model parameters with respect to each pattern are estimated by least squares estimation. The rotation invariance property, which is desired in texture classification, is maintained by applying the Fourier transform. From the experimental results, it can be concluded that: 1) the proposed method makes significant improvement over the state-of-the-art descriptors; 2) image microscopic configuration contains robust and reliable discriminative information which can improve the classification performance; and 3) the robustness of LCPs against illumination variations and image rotation is strongly implied.

References

[1] T. Ahonen, A. Hadid, and M. Pietikäinen. Face recognition with local binary patterns. In *Proc. ECCV*, volume 3021, pages 469–481, 2004.

[2] T. Ahonen, J. Matas, C. He, and M. Pietikäinen. Rotation invariant image description with local binary pattern histogram fourier features. In *Proc. SCIA*, volume 5575, pages 61–70, 2009.

[3] B. Caputo, E. Hayman, and P. Mallikarjuna. Class-specific material categorisation. In *Proc. ICCV*, pages 1597–1604, 2005.

[4] R. Chellappa and S. Chatterjeey. Classification of textures using gaussian markov random fields. *IEEE ASSP*, 33:959–963, 1985.

[5] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proc. CVPR*, volume 1, pages 886–893, 2005.

[6] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink. Reflectance and texture of real world surfaces. *ACM TG*, 18:1–34, 1999.

[7] H. Deng and D.A. Clausi. Gaussian mrf rotation-invariant features for image classification. *IEEE T-PAMI*, 26(7):951–955, 2004.

[8] Z. Guo, L. Zhang, and D. Zhang. A completed modeling of local binary pattern operator for texture classification. *TIP*, 19:1657–1663, 2010.

[9] Z. Guo, L. Zhang, and D. Zhang. Rotation invariant texture classification using lbp variance (lbpv) with global matching. *PR*, 43:706–719, 2010.

[10] Z. Guo, L. Zhang, D. Zhang, and S. Zhang. Rotation invariant texture classification using adaptive lbp with directional statistical features. In *Proc. ICIP*, pages 2677–2680, 2010.

[11] R.L. Kashyap and A. Khotanzad. A model-based method for rotation invariant texture classification. *IEEE T-PAMI*, 8(7):472–481, 1986.

[12] S. Liao and C.S. Chung. Dominant local binary patterns for texture classification. *TIP*, 18:1107–1118, 2009.

[13] D. Lowe. Object recognition from local scale-invariant features. In *Proc. ICCV*, pages 1150–1157, 1999.

[14] D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, pages 91–110, 2004.

[15] S.G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE T-PAMI*, 11:674–693, 1989.

[16] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on feature distributions. *PR*, 29(1):51–59, 1996.

[17] T. Ojala, T. Mäenpää, M Pietikäinen, J. Viertola, J. Kyllönen, and S. Huovinen. Outex-new framework for empirical evaluation of texture analysis algorithm. In *Proc. ICPR*, pages 701–706, 2002.

[18] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE T-PAMI*, 24(7):971–987, 2002.

[19] M. Pietikäinen, T. Nurmela, T. Mäenpää, and M. Turtinen. View-based recognition of real-world textures. *PR*, 37(2):313–323, 2004.

[20] T. Randen and J.H. Husoy. Filtering for texture classification: A comparative study. *IEEE T-PAMI*, 21:291–310, 1999.

[21] S. Roweis and L. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326, 2000.

[22] M. Unser. Texture classification and segmentation using wavelet frames. *TIP*, 11:1549–1560, 1995.

[23] M. Varma and A. Zisserman. Classifying images of materials: Achieving viewpoint and illumination independence. In *Proc. ECCV*, volume 2352, pages 255–271, 2002.

[24] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *IJCV*, 62:61–81, 2005.

[25] M. Varma and A. Zisserman. A statistical approach to material classification using image patch exemplars. *IEEE T-PAMI*, 31(11):2032–2047, 2009.

[26] J. Weszka, C.R. Dyer, and A. Rosenfeld. A comparative study of texture measures for terrain classification. *IEEE TSMC*, 6:269–285, 1976.

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