Pairwise Rotation Invariant Co-occurrence Local Binary Pattern

Xianbiao Qi[1,⋆], Rong Xiao[2], Jun Guo[1], and Lei Zhang[2]

[1] Beijing University of Posts and Telecommunications, Beijing 100876, P.R. China
[2] Microsoft Research Asia, Beijing 100080, P.R. China

Abstract. In this work, we introduce a novel pairwise rotation invariant co-occurrence local binary pattern (PRI-CoLBP) feature which incorporates two types of context - spatial co-occurrence and orientation co-occurrence. Different from traditional rotation invariant local features, pairwise rotation invariant co-occurrence features preserve relative angle between the orientations of individual features. The relative angle depicts the local curvature information, which is discriminative and rotation invariant. Experimental results on the CUReT, Brodatz, KTH-TIPS texture dataset, Flickr Material dataset, and Oxford 102 Flower dataset further demonstrate the superior performance of the proposed feature on texture classification, material recognition and flower recognition tasks.

1 Introduction

Designing effective features is a fundamental problem in many computer vision tasks, including image retrieval, object and scene recognition, texture classification, etc. However, in many real applications, there always exists large intra-class variation due to different 3D poses and different object appearance. Sometimes, intra-class variation is even larger than inter-class variation when the objects from different classes have the same 3D pose. It therefore becomes crucial to design discriminative features.

It has been proved that co-occurrence of features could boost the discriminative power of features. Observing co-occurrence of two features could provide much more information than observing occurrence of two features individually. Roughly, existing co-occurrence methods can be divided into two categories. One category does not consider the spatial co-occurrence, such as the work [1]. In this work, authors proposed to mine co-occurrence statistics of SIFT [2] words in a whole image for visual recognition. In contrast, the other category takes the spatial co-occurrence into account, such as the color co-occurrence histogram [3] and co-occurrence of histogram of orientation gradient(CoHoG) [4]. Color co-occurrence just calculates co-occurrence statistics of color pixels in a fixed spatial distance K, and CoHoG is a co-occurrence histogram of pairs of gradient orientations with various positional offsets.

⋆ This work was performed at Microsoft Research Asia.

A. Fitzgibbon et al. (Eds.): ECCV 2012, Part VI, LNCS 7577, pp. 158–171, 2012.
© Springer-Verlag Berlin Heidelberg 2012

However, many vision applications also suffer from the geometric and photometric variations. So, how to design invariant features becomes another important vision problem and have been studied over decades. Unfortunately, few works in the literature have investigate the problem of invariant properties of spatial co-occurrence features.

Different from traditional features defined on a single point, co-occurrence features are defined on two points, which inevitably introduces the problem of point correspondence when considering the transformation invariance. More specifically, considering a spatial co-occurrence feature defined on two points, we shall first guarantee the correct point correspondence after a transformation then study the feature invariance. We call such kind of transformation invariance "pairwise transformation invariance (PTI)". Moreover, points used to define co-occurrence feature usually have a predetermined structure, which means we can select one point as the base point and find other points based on a set of mapping functions. Using this method, a co-occurrence feature can be regarded as a traditional feature defined on a single point.

In this paper, we investigate the problem of pairwise transformation invariance, and point out that a PTI feature can be regarded as a transformation invariant feature when these mapping functions meet some constraints. Based on these studies, we extend the traditional local binary pattern (LBP) feature to a pairwise rotation invariant co-occurrence LBP feature. Instead of simply encoding co-occurrence of two individual rotation invariant feature, we propose an effective encoding strategy, which well preserves the information of relative orientation angle θ as shown in Fig. 1. The relative orientation angle θ is rotation invariant and capable of capturing local curvature information which is discriminative. Moreover, we propose to extend the proposed co-occurrence feature to encode multi-scale and multi-orientation information.

Experimental results show that the proposed PRI-CoLBP is rotation invariant and effective for many applications. Compared with the state-of-the-art approaches on texture classification, material recognitions and flower recognition on five benchmark datasets, our feature demonstrates superior performance.

The rest of the paper is organized as follows. In section 2, we first introduce some related works. And then, in section 3, we elaborate our proposed methods. Experimental results will be presented in section 4. Finally, we give a conclusion and discuss future works in section 5.

2 Preliminaries on LBP

The LBP operator was first proposed by Ojala et al. [5] as a gray-scale invariant texture measure. For each pixel on the image, threshold its circularly symmetric neighbor set of P members $N_p(x)$ on a circle of radius R with the center value and consider the result as a binary number. The LBP operator $\psi_{P,R}(x)$ could be written as follows:

$$\psi_{P,R}(x) = \sum_{p=0}^{P-1} s_p(x) 2^p \tag{1}$$

$$s_p(x) = \begin{cases} 1, & V(N_p(x)) \geq V(x) \\ 0, & V(N_p(x)) < V(x) \end{cases} \tag{2}$$

Where x is a coordinate, $N_p(x)$ is the pth neighbor of point x, and V(x) is the pixel value of point x. Note that, the function $s_p(x)$ isn't affected by changes in mean luminance, hence, LBP could achieve gray-scale invariance.

The image rotation always results in a rotation permutation of pixel R-Neighbor set $N_p(x)$ and possibly changes the LBP value. In order to achieve rotation invariance, Ojala et al.[5] defines the rotation invariant LBP as follows:

$$\psi^{ri}_{P,R}(x) = min\{ROR(\psi_{P,R}(x), i) \ \mid \ i \in [0, P-1]\} \tag{3}$$

where $ROR(x, i)$ performs a circular bit-wise right shift for i times on P-bit number x.

Ojala et al. also observed that patterns with very few spatial transitions describe the fundamental properties of image microstructure. These patterns are called "uniform patterns". For example pattern 11110000 describes a local edge, and pattern 11111111 describes a flat region or a dark spot. To formally define the uniform patterns, they introduce a uniformity measure $U(x)$,

$$U(x) = \sum_{p=1}^{P} |s_p(x) - s_{p-1}(x)| \tag{4}$$

where $s_P(x)$ is defined as $s_0(x)$. For the uniform patterns, $U(x) \leq 2$.

To make the further discussion simple, we will limit our discussion on $LBP_{8,1}$ [1] patterns and use the term LBP instead. The result can be easily extended to general LBP patterns. Here are some statistical result of $LBP_{8,1}$ patterns which will be used in later sections. There are 58 unique uniform patterns. Usually, an additional pattern is used to represent the non-uniform patterns. Therefore, the uniform LBP histogram consists of 59 patterns. For rotation invariant uniform LBP, it consists of 10 patterns.

3 Pairwise Rotation Invariant Co-occurrence Local Binary Pattern

Different from traditional local features which are defined in the neighborhood of a single point, co-occurrence features are defined on the neighborhood of two points. As shown in Fig. 1, CoLBP features can be used to describe more complex image structure. This is essentially the same as the co-occurrence matrix introduced by Haralick et al. [6].

In this section, we first introduce a new concept called pairwise transformation invariant and analyze its properties. And then, according to the proposed concept, we design a specific pairwise co-occurrence feature called pairwise rotation invariant CoLBP(**PRI-CoLBP**). Finally, we introduce some extensions to the PRI-CoLBP.

[1] $LBP_{8,1}$ means P=8 and R=1.

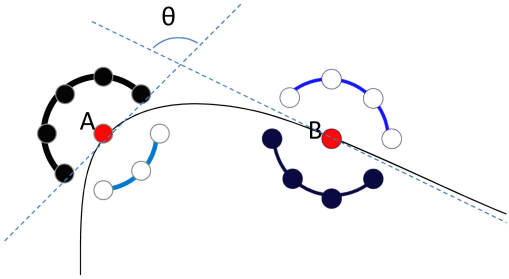

Fig. 1. An illustration of co-occurrence local binary pattern. Co-occurrence LBP could encode more complex structure. Meanwhile, no matter how the image rotates, the relative angle θ is rotation invariant, and the relative angle describes the local curvature information which is discriminative.

3.1 Pairwise Transformation Invariant

Notation: $f(A)$ is a local feature defined on the neighborhoods of point A. $F(A, B)$ is a co-occurrence feature defined on the neighborhoods of point A and point B. If point B can be determined by point A using $B = \varphi(A)$, the co-occurrence feature $F(A, B)$ can be therefore defined as $\zeta(A) = F(A, \varphi(A))$, where $\varphi(.)$ is mapping function from A to B, and $\eta(.)$ is an in-plane transformation defined on an image.

Transformation Invariant(TI): $f(A)$ is invariant to transformation $\eta(.)$ if and only if

$$f(A) = f(\eta(A)) \tag{5}$$

Pairwise Transformation Invariant(PTI): The co-occurrence feature $F(A, B)$ is pairwise invariant to transformation $\eta(.)$ if and only if

$$F(A, B) = F(\eta(A), \eta(B)) \tag{6}$$

TI Versus PTI: PTI feature is different from TI feature. PTI feature is defined on a pair of points and TI feature is defined on a single point. We can easily verify that the co-occurrence of two TI features is a PTI feature because

$$\begin{aligned} F(\eta(A), \eta(B)) &= [f_1(\eta(A)), f_2(\eta(B))]_{co} \\ &= [f_1(A), f_2(B)]_{co} \\ &= F(A, B) \end{aligned}$$

where $f_1(.)$ and $f_2(.)$ are TI features invariant to transformation $\eta(.)$, and $[f_1(\eta(A)), f_2(\eta(B))]_{co}$ can be considered as the concatenation of the two features. With different choices of local features $f_1(.)$ and $f_2(.)$, it will lead to different encoding strategies for co-occurrence features. However, pairwise transformation invariant isn't the simple combination of two transformation invariant features. From the Fig. 1, the relative orientation angle θ is rotation invariant, but the angle θ cannot be encoded by two individual rotation invariant features.

Moreover, a PTI feature is usually not a TI feature, because PTI feature is defined on two points and doesn' t guarantee that the correspondence of two points is invariant to the transformation. Fortunately, most spatial co-occurrence features have a predetermined structure, which means the two points used to define the co-occurrence feature are interdependent. Using this assumption, we can convert a PTI feature to a TI feature. We have the following lemma.

Lemma 1. Suppose $B = \varphi(A)$, then $F(A, B)$ can be denoted as $F(A, \varphi(A))$ which is defined on a single point, here, we denote $\zeta(A) = F(A, \varphi(A))$, feature $\zeta(A)$ is transformation invariant if the function F is pairwise transformation invariant and $\eta(\varphi(A)) = \varphi(\eta(A))$.

Proof:

$$\zeta(\eta(A)) = F(\eta(A), \varphi(\eta(A)) = F(\eta(A), \eta(\varphi(A))$$

since function F is PTI, using Eq. 6, we have

$$\zeta(\eta(A)) = F(A, \varphi(A)) = \zeta(A)$$

Based on Eq. 5, $\zeta(A)$ is therefore transformation invariant.

A direct explanation of Lemma 1 is that $B = \varphi(A)$ and $\eta(\varphi(A)) = \varphi(\eta(A))$ means that for reference point A, target point B could be uniquely determined. Meanwhile, the PTI property of function F guarantees that for the same pairs A and B, PTI feature $F(A, B)$ is transformation invariant.

3.2 Pairwise Rotation Invariant CoLBP

Based on the definition of co-occurrence in section 3.1, it is straightforward to derive a CoLBP feature, $\Psi^{UU}(A, B) = [\psi^U(A), \psi^U(B)]_{co}$, where $\psi^U(.)$ is a uniform LBP feature, denoted as UU-CoLBP. However, UU-CoLBP has two problems. First, its dimension is high. The dimension of UU-CoLBP is 59*59=3481. Second, it is easy to verify that UU-CoLBP is not pairwise rotation invariant.

To overcome the drawbacks of UU-CoLBP, an intuitive choice is co-occurrence of two rotation invariant features, such as $\Psi^{RU}(A, B) = [\psi^{RU}(A), \psi^{RU}(B)]_{co}$, where $\psi^{RU}(.)$ is uniform rotation invariant LBP feature. Here, we denote it as RURU-CoLBP. However, RURU-CoLBP discards the relative angle information between A and B, which captures the curvature information.

To preserve the discriminative curvature information, we propose an effective pairwise rotation invariant CoLBP(PRI-CoLBP). Before introducing it, we will first introduce some related concept and then define PRI-CoLBP.

Uniform LBP depends on the direction which determines the start point of its circular binary sequence. Defined on different directions, uniform LBP has different values. Therefore, we introduce a uniform oriented-LBP. Suppose $\eta(., \alpha)$ is a 2D plane rotation transformation with degree α. Uniform oriented-LBP could be defined as:

$$\hat{\psi}^U(A, \alpha) = \psi^U(\eta(A, -\alpha)) \tag{7}$$

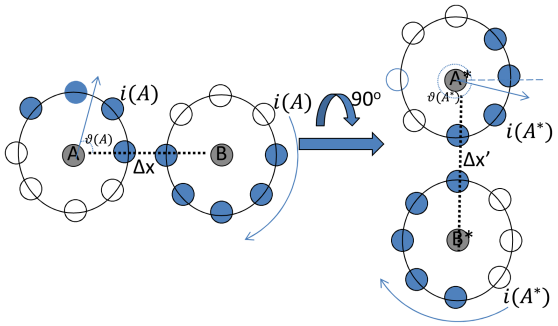

Fig. 2. An illustration of pairwise rotation invariant. For the left pairs, we first determine orientation $i(A)$ of the reference point A, then we compute the uniform pattern of B according to $i(A)$, we get co-patterns $[(00001111)_{RU}, (01111100)_U]_{co}$. For the rotated pairs in the right side, we could also get the same co-patterns.

which is defined on point A and the direction α. We can easily find that $\hat{\psi}^U(\eta(A,\beta),\alpha) = \psi^U(\eta(A,\beta-\alpha))$.

Defined on a arbitrary direction, all the neighbors are not on the grid points, it will bring in more interpolation calculations to compute the binary sequence. In practice , $\hat{\psi}^U(A,\alpha)$ can be approximated by $ROR(\psi^U(A), [\frac{8*\alpha}{2\pi}])$, where $ROR(.,\ .)$ performs a circular bit-wise right shift like Eq. 3.

Using this definition, the PRI-CoLBP feature can be defined as following:

$$\Psi^{PRI}(A,B) = [\psi^{RU}(A), \hat{\psi}^U(B,\vartheta(A))]_{co} \tag{8}$$

where $\vartheta(A)$ is an orientation determined by A, such as the gradient orientation of point A. We can easily verify that

$\Psi^{PRI}(\eta(A,\alpha),\eta(B,\alpha)) = [\psi^{RU}(\eta(A,\alpha)), \hat{\psi}^U(\eta(B,\alpha),\vartheta(\eta(A,\alpha)))]_{co}$

Note that, the first term is invariant to the rotation transformation, and using the definition of oriented-LBP, the second term can be rewritten as:

$$\hat{\psi}^U(\eta(B,\alpha),\vartheta(\eta(A,\alpha))) = \psi^U(\eta(B,\alpha-\vartheta(\eta(A,\alpha)))$$

Actually, $\alpha - \vartheta(\eta(A,\alpha)) = -\vartheta(A)$, so we have

$$\hat{\psi}^U(\eta(B,\alpha),\vartheta(\eta(A,\alpha))) = \psi^U(\eta(B,-\vartheta(A))) = \hat{\psi}^U(B,\vartheta(A))$$

Therefore the CoLBP pattern $\Psi^{PRI}(A,B)$ is pairwise rotation invariant. The function $\Psi^{PRI}(A,B)$ defined above can well capture the relative orientation information. Moreover, it is pairwise rotation invariant. As shown in Fig. 2 shows, for the left point pair, we first compute the $[\frac{8*\vartheta(A)}{2\pi}]$ and denote it as $i(A)$. Using $i(A)$ as the start point for the binary sequence of point B, we get the co-occurrence feature $[(00001111)_{RU}, (01111100)_U]_{co}$, where $(00001111)_{RU}$ means a rotation invariant uniform LBP code and $(01111100)_U$ means a uniform LBP code. For the co-occurrence pattern, we use the gradient magnitude

of two points to weight the co-pattern. For a whole image region, dense sampling of each pixels is applied to obtain an accumulated co-occurrence histogram. Since $(00001111)_{RU} = 4$ and $(01111100)_U = 8$, the accumulated histogram $H(4,8) = H(4,8) + M(A) + M(B)$, where $M(A)$ and $M(B)$ are the gradient magnitude of point A and B and H is 2D histogram with dimension 10×59.

In Eq. 8, we have introduced an effective pairwise rotation invariant function $\Psi^{PRI}(A, B)$. In order to achieve rotation invariant, we should ensure that $\eta(\varphi(.)) = \varphi(\eta(.))$ is satisfied, which means that the target point B could be uniquely ascertained no matter how the image is rotated. Denote the vector of unit gradient direction of point A as $g(A)$ and the vector of unit normal direction of point A as $n(A)$, here, we can define $\varphi(A)$ as:

$$\varphi(A) = a * g(A) + b * n(A) + A \tag{9}$$

where a and b are two parameters. A is the coordinates. $\varphi(A)$ is defined on the local polar coordinate with x axis $g(A)$ and y axis n(A). It is easy to prove that the definition of $\varphi(.)$ satisfies the condition $\eta(\varphi(.)) = \varphi(\eta(.))$. With different choices of a and b, it will lead to different function $\varphi(.)$. For example, $a = 2$ and $b = 0$ mean that the point B can be determined as 2 pixels distance from A along with the gradient direction.

For applications, which do not have large rotation variation, we can directly define the function $\varphi(.)$ along with the horizontal and vertical unit direction. In this way, the feature is non-rotation invariant, but it takes less time to process an image because for rotation invariant, there are more interpolation computations than non-rotation invariant. Therefore, in this paper, we define the rotation invariant PRI-CoLBP feature which is computed along with gradient and normal unit direction as PRI-CoLBP$_g$ and define the non-rotation invariant PRI-CoLBP feature which is computed directly along with the horizontal and vertical unit direction as PRI-CoLBP$_0$.

3.3 Extension Using Multi-scale and Multi-orientation, Color, PCA

A PRI-CoLBP co-pattern only captures local texture co-occurrence from a single scale and orientation. In order to describe multi-scale and multi-orientation local image structure, we extend PRI-CoLBP to multi-scale and multi-orientation PRI-CoLBP (MSMO PRI-CoLBP). As shown in Fig. 3, for each reference point A, we extract PRI-CoLBP patterns from point pairs (A, B_i) in multi-scale and multi-orientation. In order to achieve rotation invariance, we assume point $B_i = \varphi_i(A)$ and the function $\varphi_i(.)$ is defined on the local polar axis with the unit gradient direction as its x axis and the unit normal direction as its y axis. Using this strategy, we extract 6 co-occurrence patterns in three scales and two orientations, resulting in a vector with dimension $6 \times 590 = 3540$.

For applications, such as flower recognition [7][8][9], color has proven to be complementary to texture and shape information. In order to incorporate color information into PRI-CoLBP feature, MSMO PRI-CoLBP feature in three color channels are extracted and concatenated into a vector with dimension $3540 \times 3 = 10620$. Fortunately, many patterns don't co-occur together. Thus, the resulting

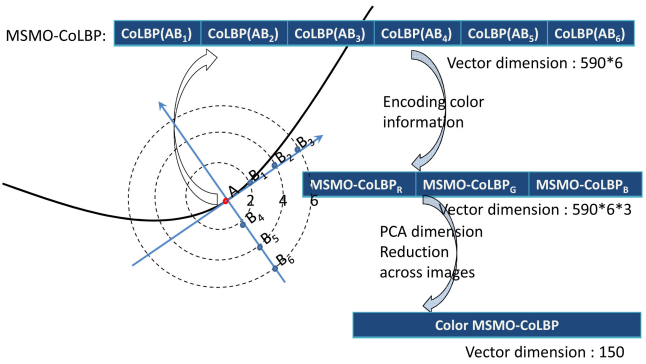

Fig. 3. MSMO PRI-CoLBP features are extracted in the local polar coordinate system defined on gradient direction and normal direction. First, PRI-CoLBP features from two scales and three orientations are concatenated. Then, multi-scale multi-orientation CoLBP features from three color channels are extracted and concatenated. Finally, PCA algorithm is used to remove redundancy.

co-occurrence feature is very sparse and standard dimensionality reduction technologies can be applied, such as principal component analysis(PCA). Based on our empirical study, when we decrease the dimension of PRI-CoLBP to around 120 to 150, it gets similar and even better performance than the original PRI-CoLBP.

4 Experimenatal Results

In this section, we first introduce some experimental details. Then, we compare three kinds of different encoding strategies. After that, we validate the rotation invariant property of PRI-CoLBP. Finally, we show the applications of PRI-CoLBP on texture classification, material and flower recognition.

4.1 Implementation Details

In this paper, we evaluate the proposed feature on Brodatz[10], CUReT[11], KTH-TIPS[12], Flickr Material Database(FMD) [13] and Oxford Flower 102 data set [7]. For all datasets, we extract MSMO PRI-CoLBP features in 3 scales and 2 orientations. Since we extract MSMO PRI-CoLBP in all experiments, to make the notation simple, we abbreviate MSMO PRI-CoLBP as PRI-CoLBP. For Brodatz, KTH-TIPS and FMD, only gray scale images are used. For CUReT and Oxford Flower 102, 3 color channels are utilized to extract color PRI-CoLBP' features. For the three texture data sets, $LBP_{8,2}$ is used to calculate PRI-CoLBP. For flower and FMD, $LBP_{8,1}$ is used instead.

Classification Method. SVM classifier with χ^2 kernel is widely used and proven effective in [14]. In [15], Vedaldi et al. propose an efficient additive kernel approximation(AKA) method, which enables fast training/evalution using

linear SVM and is scalable to large data set. When PCA is applied to remove redundancy, SVM with RBF kernel will be used. For all SVM classifiers, one-vs-all scheme is used. We will compare these three classification strategies. For fair comparison with other features, we use the same classification method.

Notation. As described in section 3.2, we define two kind of PRI-CoLBP, non-rotation invariant PRI-CoLBP$_0$ and rotation invariant PRI-CoLBP$_g$. For three texture classification tasks, there is only small rotation variation in these datasets. Therefore, we use PRI-CoLBP$_0$. For FMD and flower recognition, we use PRI-CoLBP$_g$. In section 4.3, we will compare PRI-CoLBP$_0$ and PRI-CoLBP$_g$.

Computational Cost. Like the traditional LBP operator, the extraction for PRI-CoLBP is very fast. On a desktop computer with dual-core 2.8G CPU, the c++ implementation of PRI-CoLBP$_0$ takes about 0.005 second for an 200×200 image and it takes about 0.031 second to extract PRI-CoLBP$_g$. The fast extraction property enables the proposed feature for large scale problem.

4.2 Encoding Strategies for CoLBP

In section 3.2, we propose three CoLBP features, PRI-CoLBP, RURU-CoLBP and UU-CoLBP, where PRI-CoLBP and RURU-CoLBP are pairwise rotation invariant and UU-CoLBP is not. We conduct experiments to compare the effectiveness of these three encoding strategies. For all experiments, we use direct χ^2 kernel SVM. For Brodatz, KTH-TIPS and FMD, we separately use 3, 40 and 50 samples for training and the rest for test. The results are shown in Tab. 1.

Table 1. Comparison of encoding strategies: RURU-CoLBP, UU-CoLBP, PRI-CoLBP

Dataset	RURU-CoLBP	UU-CoLBP	PRI-CoLBP
Brodatz(3)	93.9%	95.7%	**96.6%**
KTH-TIPS(40)	96.4%	97.9%	**98.3%**
FMD(50)	52.1%	50.6%	**56.5%**

From Table 1, we can find that PRI-CoLBP outperforms the other two encoding strategies in all three data sets. Especially on FMD data set, PRI-CoLBP achieves about 5% higher performance. It means that the relative orientation angle in Fig. 1 is discriminative and PRI-CoLBP is effective on capturing the curvature information.

The reason why UU-CoLBP works well on Brodatz and KTH-TIPS is that these two data sets don't have large rotation variation. But on FMD data set with normal rotation variation, UU-CoLBP works worse than RURU-CoLBP. Overall, PRI-CoLBP performs well on all datasets, it proves the effectiveness of the spatial co-occurrence and the relative angle information is well described.

4.3 Rotation Invariant

In this subsection, we conduct experiments to validate the rotation invariance property of PRI-CoLBP$_g$. Here, we will compare PRI-CoLBP$_g$ and PRI-CoLBP$_0$

in the cases of with and without notable rotation variance. The experiments are conducted on KTH-TIPS, CUReT and FMD data sets. As there are only small rotation variation in KTH-TIPS and CUReT, we generate two new data-sets by adding arbitrary rotations to the two data-sets.

For the images int the FMD data-set have normal rotation variations, so we just compare PRI-CoLBP$_0$ and PRI-CoLBP$_g$ directly. For KTH-TIPS and CUReT, we perform two experimental comparisons on original data-sets and the newly generated data-sets with arbitrary rotations.

Newly Generated Data-sets: For KTH-TIPS, we first rescale each image to 1.5 times of its original size and then randomly rotate the image to a random angle. After that, we crop an image of same size as the original image from the center of the rotated image. This can guarantee that no black region will be cropped out. For CUReT, each image is rotated randomly and then a sub image of size 141*141 is cropped out from the center of the rotated images. In this way, we can create a corresponding modified data-set with same number of images. Note that, the rotating and cropping will lead to information loss and rotation variations. Same training samples and classification methods are used. The experimental results are shown in Tab. 2.

Table 2. Comparison of PRI-CoLBP$_0$ and PRI-CoLBP$_g$ on KTH-TIPS, CUReT, FMD data sets and modified KTH-TIPS and CUReT data sets

	Rotation	PRI-CoLBP$_0$	PRI-CoLBP$_g$
KTH-TIPS(40)	No Rotation	**98.8%**	**98.8%**
	Arbitrary Rotation	90.8%	**96.4%**
CUReT(46)	No Rotation	**99.2%**	98.9%
	Arbitrary Rotation	89.1%	**94.7%**
FMD(50)	Normal Rotation	53.4%	**56.5%**

From Tab. 2, we can see that PRI-CoLBP$_g$ outperforms PRI-CoLBP$_0$ for 5.6% and 5.6% on the modified KTH-TIPS and CUReT, which have arbitrary rotation. It proves the effectiveness of the rotation invariant property of PRI-CoLBP$_g$. On FMD with normal rotation, PRI-CoLBP$_g$ also outperforms PRI-CoLBP$_0$ for 3.1%, It further proves the effectiveness of rotation invariance. On the datasets with small rotation variations, both PRI-CoLBP$_g$ and PRI-CoLBP$_0$ work well and have similar performance. We can also find that the average performance on the modified data sets is lower than the performance on the original data set. This is mainly due to the heavy information loss and rotation variations brought by rotating and cropping.

4.4 Applications

Texture Classification: Brodatz contains 111 classes, each with 9 images. In our experiments, we use 3 samples for training and the rest for testing.

For CUReT, we use the same subset of images as [16][17], which contains 61 texture classes, each with 92 images. We use 46 samplers per class for training and the rest 46 for testing. For KTH-TIPS, it consists of 10 texture classes, each with 81 images, captured at nine scales, viewed under three different illumination directions and three different poses. We use 40 images per class for training and the rest for testing. The work of Zhang et al. [14] is widely recognized as the state-of-the-art method. Recently, Nguyen et al. [18] proposed to use multivariate log-gaussian cox processes for texture classification.

The experimental results are shown in Tab. 3. All the results are originally reported except the result of multi-scale LBP(MSLBP) which is based on the standard implement[2] using SVM with χ^2 kernel.

Table 3. Texture Classification results on Brodatz, CUReT and KTH-TIPS data-sets. For Brodatz, 3 training images are used and 46 for CUReT, 40 for KTH-TIPS.

Methods	Brodatz(3)	CUReT(46)	KTH-TIPS(40)
PRI-CoLBP$_0$	**96.6%**	98.6%	98.3%
PRI-CoLBP$_0$(PCA)	96.0%	**99.2%**	**98.8%**
PRI-CoLBP$_0$(AKA)	**96.8%**	98.5%	97.6%
Nguyen et al.[18]	96.1%	—	95.7%
Zhang et al.[14]	95.9%	95.3%	96.1%
VZ-Patch[16]	92.9%	98.0%	92.4%
Caputo et al.[17]	95.0%	98.5%	94.8%
Lazebnik et al.[19]	88.2%	72.5%	91.3%
MSLBP[5]	91.6%	96.3%	92.2%

From Tab. 3, PRI-CoLBP$_0$ significantly outperforms its direct competitor MSLBP on all the three datasets. Meanwhile, PRI-CoLBP$_0$ obviously outperforms [14] and [18] on all the three datasets. Compared with [14] and [18], the proposed feature reduces the average error rate from about 4% to 2% which is equivalent to reduce the errors by 50%. Although all classification methods have slightly different performance, they generally work well.

The reason why PRI-CoLBP$_0$ greatly outperforms MSLBP is that PRI-CoLBP$_0$ effectively captures spatial context co-occurrence and preserves relative orientation angle information, and the spatial context and relative orientation information are greatly discriminative, whereas multi-scale LBP ignores the spatial information. Although KTH-TIPS contains scale variation (9 scales and 81 images per class), the scale changes continually from 0.5 to 2 and the variation is relatively small. Thus, PRI-CoLBP$_0$ works well on this data set and obtains an accuracy of 98.3%, which significantly outperforms the best result 96.1%. The experimental result on KTH-TIPS indirectly reflects that PRI-CoLBP$_0$ is insensitive to small scale variation.

[2] `http://www.cse.oulu.fi/MVG/Downloads`

Material Recognition: FMD is a newly published material dataset, which is considered to be challenging. FMD contains 10 classes, each with 100 images. We use 50 samples for training and the rest for test. Experimental results are shown in Tab. 4, where MSLBP is obtained using standard implementations.

Table 4. Experimental results on FMD dataset using 50 training samples

Methods		50
Single Feature	SIFT [21]	35%
	MSLBP	43.5%
	Kernel Descriptor in [20]	49%
	PRI-CoLBP$_g$	**56.5%±1.8**
	PRI-CoLBP$_g$(PCA)	**55.2%±1.8**
	PRI-CoLBP$_g$(AKA)	**57.4%±1.7**
Multiple Features	Liu et al. CVPR 2010 [21]	44.6%
	Hu et al. BMVC 2011 [20]	54% ±2.0

From Tab. 4, for single feature, PRI-CoLBP$_g$ outperforms MSLBP for about 13% and bag-of-sift [21] for about 23%. In addition, it outperforms the best single feature for 7.5% from 49% to 56.5%. Moreover, using same kernel SVM classifier, PRI-CoLBP$_g$ also outperforms the previously published best performance 54% [20] which combines five kernel descriptors.

Material recognition is highly related to texture classification. Although FMD has normal rotation variance and the recognition task is challenging because the material exhibits different appearances, PRI-CoLBP$_g$ is highly effective in preserving relative angle information between co-occurrence pairs as shown in Sec. 4.2 and also rotation invariant as shown in Sec. 4.3. As a result, PRI-CoLBP$_g$ works very well on FMD.

Flower Recognition: Oxford Flowers 102 [7] has 8189 images which are divided into 102 categories with 40 to 250 images. In [7][8][9], the authors prove that classification with segmentation can boost the final performance. Instead of focusing on segmentation, we use Grabcut [22] to segment the images. We resize the foreground images to a minimum resolution of 128 and extract color PRI-CoLBP$_g$ feature the foreground region.

The work of [7] is widely recognized as the state-of-the-art method. Recently, Chai et al [9] achieved better result. We use the same training and testing methods as in [9] and use 20 and 30 samples for training and the rest for test.

According to [9], we reimplement the bag of Multi-Scale Dense Sampling (MSDS) SIFT features and get performance 69.5%, which is close to their reported result 70.0%. Experimental result are shown in Tab. 5.

From Tab. 5, for single feature, PRI-CoLBP$_g$ outperforms MSLBP for 27% and another co-occurrence feature(COHED) for more than 30%. In addition, PRI-CoLBP$_g$ also outperforms the best two single feature methods up to date.

Table 5. Recognition performance for single features and multiple features on Oxford Flower 102 dataset

Single Features	20	30	Multiple Features	20	30
SIFT boundary [7]	32.0%	–	Ito et al. [4]	53.2%	–
HSV [7]	43.0%	–	Nilsback et al. [7]	72.8%	–
HOG [7]	49.6%	–	Yuan et al. [23]	74.1%	–
SIFT internal [7]	55.1%	–	Nilsback's thesis [8]	76.8%	–
CoHED [4]	48.2%	–	Grabcut [9]	77.0%	–
MSLBP	52.0%	–	Chai et al.[9]	80.0%	–
MSDS [9]	69.5%	73.4%	PRI-CoLBP$_g$ + MSDS	**84.2%**	**87.1%**
Kanan et al [24]	71.4%	75.2%			
PRI-CoLBP$_g$	**79.1%**	**82.3%**			

Specifically, it outperforms MSDS for about 10% and outperforms the work [24] for 7.7%. For multiple features combinational methods, PRI-CoLBP$_g$ combined with MSDS achieves an accuracy of 84.2%, which outperforms nowadays published best result 80.0% [9] which combines four types of features.

The reason why PRI-CoLBP$_g$ works so well on flower recognition is that flower recognition is highly related to color texture classification and color PRI-CoLBP$_g$ well captures color and texture information. PRI-CoLBP$_g$ combined with MSDS greatly outperforms each of single feature, it means that PRI-CoLBP$_g$ is complementary to the MSDS.

5 Conclusion and Future Works

In this paper, we have studied the problem of pairwise transformation invariance, and point out that a PTI feature can be regarded as a transformation invariant feature when mapping functions meet some constraints. Based on the studies, we propose an effective pairwise rotation invariant CoLBP feature which aims to preserve relative angle information between co-occurrence pairs. The experimental results show that the proposed feature effectively captures local curvature information and demonstrates a great rotation invariant property. Moreover, the computational cost for the proposed feature is low since most computation comes from calculation of LBP and the image gradient.

In our future work, we will apply the proposed feature on more applications, such as food recognition, leaf recognition, and large scale scene classification, and test its generalization ability on these applications.

Acknowledgements. This work was supported by National Natural Science Foundation of China(Grant No.61005004 and 61175011),the 111 project(Grant No.B08004), and the Fundamental Research Funds for the Central Universities(Grant No.2012RC0108).

References

1. Yuan, J., Yang, M., Wu, Y.: Mining discriminative co-occurrence patterns for visual recognition. In: CVPR (2011)
2. Lowe, D.: Distinctive image features from scale-invariant keypoints. IJCV (2004)
3. Chang, P., Krumm, J.: Object recognition with color co-occurrence histograms. In: CVPR (1999)
4. Ito, S., Kubota, S.: Object Classification Using Heterogeneous Co-occurrence Features. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part II. LNCS, vol. 6312, pp. 209–222. Springer, Heidelberg (2010)
5. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. PAMI (2002)
6. Haralick, R., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics (1973)
7. Nilsback, M., Zisserman, A.: Automated flower classification over a large number of classes. In: ICVGIP (2008)
8. Nilsback, M.: An automatic visual flora - segmentation and classification of flowers images. PhD thesis, University of Oxford (2009)
9. Chai, Y., Lempitsky, V., Zisserman, A.: Bicos: A bi-level co-segmentation method for image classification. In: ICCV (2011)
10. Brodatz, P.: Textures: a photographic album for artists and designers. Dover Publications, New York (1999)
11. Dana, K., Van Ginneken, B., Nayar, S., Koenderink, J.: Reflectance and texture of real-world surfaces. ACM Transactions on Graphics (TOG) (1999)
12. Hayman, E., Caputo, B., Fritz, M., Eklundh, J.-O.: On the Significance of Real-World Conditions for Material Classification. In: Pajdla, T., Matas, J. (eds.) ECCV 2004, Part IV. LNCS, vol. 3024, pp. 253–266. Springer, Heidelberg (2004)
13. Sharan, L., Rosenholtz, R., Adelson, E.: Material perception: What can you see in a brief glance? Journal of Vision (2009)
14. Zhang, J., Marszalek, M., Lazebnik, S., Schmid, C.: Local features and kernels for classification of texture and object categories: a comprehensive study. In: CVPR (2007)
15. Vedaldi, A., Zisserman, A.: Efficient additive kernels via explicit feature maps. In: CVPR (2010)
16. Varma, M., Zisserman, A.: A statistical approach to material classification using image patch exemplars. PAMI (2008)
17. Caputo, B., Hayman, E., Fritz, M., Eklundh, J.: Classifying materials in the real world. Image and Vision Computing (2010)
18. Nguyen, H., Fablet, R., Boucher, J.: Visual textures as realizations of multivariate log-gaussian cox processes. In: CVPR (2011)
19. Lazebnik, S., Schmid, C., Ponce, J.: A sparse texture representation using local affine regions. PAMI (2005)
20. Hu, D., Bo, L.: Toward robust material recognition for everyday objects. In: BMVC (2011)
21. Liu, C., Sharan, L., Adelson, E., Rosenholtz, R.: Exploring features in a bayesian framework for material recognition. In: CVPR (2010)
22. Rother, C., Kolmogorov, V., Blake, A.: Grabcut: Interactive foreground extraction using iterated graph cuts. ACM Transactions on Graphics (TOG) (2004)
23. Yuan, X., Yan, S.: Visual classification with multi-task joint sparse representation. In: CVPR (2010)
24. Kanan, C., Cottrell, G.: Robust classification of objects, faces, and flowers using natural image statistics. In: CVPR (2010)