LOCAL BINARY PATTERN HISTOGRAM BASED TEXTON LEARNING FOR TEXTURE CLASSIFICATION

Yonggang He, Nong Sang and Rui Huang

Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan, 430074, P.R. China

ABSTRACT

Local Binary Pattern (LBP) and Texton are both widely used texture analysis techniques. In this paper we propose a patch-based texture classification method that takes advantage of both LBP and Texton. Unlike the traditional LBP methods that describe a texture with the occurrence of local binary patterns in the entire image, we compute the LBP histogram in a small region around each pixel to capture the local structure information. The texton learning method is then performed on these LBP histograms, resulting in a texture classification algorithm that outperforms the traditional LBP-based methods due to its preservation of local structure information. It also outperforms the traditional filtering-based texton methods due to its robustness to orientation and illumination. Experimental results on two benchmark databases validate the advantages of the proposed method.

Index Terms— texture classification, local binary pattern, texton

1. INTRODUCTION

Texture classification as an important research topic in image processing and pattern recognition has been extensively studied during the last several decades. The early representative texture classification methods are the co-occurrence matrix method[1] and filtering-based approaches[2, 3]. These methods are mostly sensitive to orientation and illumination changes. Recently, texton-based methods have shown their power in texture classification. Leung and Malik[4] obtained good texture classification results by building 3D textons from filter responses over a stack of texture images under different viewpoints and illuminations. Cula and Dana[5] generated 3D textons from single texture images instead of image stacks, although their method is not rotation invariant. Varma and Zisserman[6] proposed the MR8 algorithm, which creates rotation invariant textons from 8 maximum filter responses. They later achieved even better performances[7] by replacing the filter responses with the image patch appearance.

Another category of powerful texture classification algorithms use the Local Binary Pattern (LBP) operator, which is based on the signs of differences of neighboring pixels. Since the LBP was proposed by Ojala et al.[8], it has been applied to various texture analysis tasks[9, 10]. Ojala et al. showed the ability of the LBP operator for texture classification in [11]. They later proposed the Rotation Invariant Uniform Pattern (LBP_{RIU}) [12] to achieve rotation invariance. A more recent development of this line of work is the LBP histogram Fourier features (LBP-HF) proposed by Ahonen et al.[13].

The above two types of texture classification methods in fact have a close relationship. Most previous LBP-based methods compute the occurrence of all kinds of local binary patterns in the entire texture image to represent the texture. Consider the LBP operator a non-parametric filtering process, and each unique LBP code a texton, the traditional LBP-based methods become a special case of the filtering-based texton learning methods. Apparently by using such “micro textons”, the LBP methods lost a lot of local structure information that can be captured by the regular filtering processes used by the other texton learning methods. For instance, as shown in Fig. 1, two small regions in the same texture image may be quite different and provide different structure information about the texture. The LBP based methods often ignore the salient local structure differences by extracting the LBP histogram in a large region or even the entire image. On the other hand, we do appreciate the LBP operator’s robustness to illumination and rotation. So in this paper we investigate...
the possibility of combining the LBP features with the texton learning method. We propose to use the LBP histogram of a small region around each pixel in the texture image to replace the filter responses originally used by the texton learning. In particular, we use $LBP_{P,R}^{nu2}$ histograms to describe the information of small regions of the texture images. And the LBP textons that represent the textures are learned by clustering the $LBP_{P,R}^{nu2}$ histograms. Once the LBP texton dictionary has been built, each pixel of a new texture image is labeled with a texton according to the similarity between the LBP histogram of the small region around that pixel and the learned textons, and the texture can be described with the occurrence of the LBP textons. The performance of the proposed approach is evaluated on the Outex database and the CUReT database.

2. LBP HISTOGRAM TEXTONS FOR TEXTURE CLASSIFICATION

In this section, we first briefly review the LBP features. The texton learning is then introduced to build the texton dictionary upon the LBP histograms. Finally, the texture description and classification method is presented.

2.1. Local binary patterns

The LBP operator[11] labels the image pixel a local pattern which is computed by comparing its gray value with those of its neighbors. Considering $P$ equally spaced sampling points on a circular neighborhood with radius $R$, the local binary pattern of the central pixel is defined as:

$$LBP_{P,R}(x, y) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p$$  \hspace{1cm} (1)

where $s(x)$ is a function used to create the binary values. To achieve the rotation invariance, Ojala et al.[12] proposes the rotation invariant uniform patterns ($LBP_{P,R}^{nu2}$) by removing the weights and limiting the transition (0/1 or 1/0 changes) of the LBP code. The rotation invariant uniform pattern of a pixel is defined as:

$$LBP_{P,R}^{nu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c), & \text{if } U(LBP_{P,R}) \leq 2 \\ P + 1, & \text{otherwise} \end{cases}$$  \hspace{1cm} (3)

where

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)|$$  \hspace{1cm} (4)

According to the definition of Equ. (3), the $LBP_{P,R}^{nu2}$ operator has $P+2$ distinct output values. Combining with the local variant ($VAR_{P,R}$), the $LBP_{P,R}^{nu2}VAR_{P,R}$ operator often gets better performance in texture classification.

2.2. LBP Histogram Texton Learning

The rotation invariant local binary pattern ($LBP_{P,R}^{nu2}$) of every pixel in the image is extracted to build the LBP image. The conventional LBP methods represent the texture image by estimating the distribution of local binary patterns of the entire image. Although such description performs well in many texture classification tasks, it has the drawback of losing local statistical information of the texture. Fig. 1 shows the $LBP_{P,R}^{nu2}$ histograms of two small regions of a texture in Outex database. It is clear that the $LBP_{P,1}^{nu2}$ histograms of the two regions are very different, because they stand for two different kinds of structure information of the texture. The conventional LBP methods ignore the difference of these small regions, because they computed the distribution of patterns on the whole image.

At this point, we propose to compute the histogram of the LBP code in a small region around each pixel. Moreover, multi-resolution technique[12] that combines the LBP histograms in different sampling radii is also employed to enhance the classification results. Hence, each class of texture images has a set of $LBP_{P,R}^{nu2}$ histograms of small regions. And every element in the set describes the local statistical information of the textures. The textons of each texture class are learned from the histogram set via the classical K-means clustering method. The textons from all the classes form the texton dictionary.

2.3. Texture description and classification

Once the LBP histogram texton dictionary has been built, a new texture can be described as the occurrence of the textons (i.e., a texton histogram). For a texture image, each pixel is labeled with the element of the texton dictionary that is closest to the $LBP_{P,R}^{nu2}$ histogram of the small region around that pixel. A texton histogram is formed by estimating the frequencies of texton labels of the texture image. In the training stage, a texton histogram is built for each gallery image.

In the classification phase, a texton histogram is built for an input query image, and the dissimilarity of the query and the gallery histograms can be evaluated as a test of goodness-of-fit. The chi-squares statistic as one of the goodness-of-fit metrics shows its good performance in texture classification [6, 7]. The chi-square distance between a query texton histogram $S$ and a gallery $M$ is computed as follows:

$$D(S, M) = \sum_{b=1}^{B} \frac{(S(b) - M(b))^2}{S(b) + M(b)}$$  \hspace{1cm} (5)
where $B$ is the number of bins and $S(b)$ and $M(b)$ are, respectively, the values of the query and the gallery at the $b$th bin.

3. EXPERIMENTS

The proposed method was tested on two public texture databases used by many literatures [4, 12, 6, 7, 13]: Outex and CUReT, whose texture images were collected under varied conditions (source of illumination, orientation and angle of view). To show the performance of our method, five state-of-the-art algorithms (VZ MR8[6], VZ_Patch[7], LBPriu2, LBPriu2/VAR, LBP-HF [12] and LBP-HF[13]) were compared to. For the texton learning methods (VZ_MR8, VZ_Patch and ours), 2000 features were randomly selected from each training sample for learning and 20 textons were learned for each texture class. Due to the sensitivity during the K-means clustering, each experiment was executed 10 times with random initializations, and the average results were reported. For the LBP-based methods (LBPriu2, LBPriu2/VAR, LBP-HF and ours), multi-resolution technique[12, 13] was used to enhance the performance. Therefore, the multi-resolution $LBP_{riu2}$ histogram was a 54 ($54=10+18+26$, according to Equ. 3) dimensional vector. We empirically choose a circular region with radius of 3 pixels.

3.1. Experiment on Outex database

The Outex database contains 24 classes of textures that are captured under three different illuminants (‘inca’, ‘t184’ and ‘horizon’) and nine different rotational angles ($0^\circ$, $5^\circ$, $10^\circ$, $15^\circ$, $30^\circ$, $45^\circ$, $60^\circ$, $75^\circ$ and $90^\circ$). And each setting contains 20 non-overlapping 128×128 texture samples for each class. Our experiments were performed on two public test suites of Outex: Outex_Tc_00010 (TC10) and Outex_Tc_00012 (TC12). The TC10 was used for studying rotation invariant texture classification, while TC12 was used for studying rotation and illumination invariant texture classification. The experimental setups of TC10 and TC12 were: 1. The TC10 test suite contained 4320 texture images under illumination ‘inca’. Each texture class had 180 samples under 9 angles (20 samples every rotational angles). For every texture class, the classifiers were trained with the samples of angle $0^\circ$, while the other 160 (20×8) samples were used for testing the classifier. Hence, there were 480 (24×20×1) models and 3840 (24×20×8) testing samples in total. 2. The TC12 test suite contained 9120 texture images in total. The classifiers were trained with the same texture images (20 samples of illuminant ‘inca’ and angle $0^\circ$ in each texture class) in TC10 test suite. And all samples that were captured under illuminant ‘t184’ or ‘horizon’ were used for testing. Hence, there were 480 (24×20×1) models and 8640 (24×20×8×2) validation samples in total.

<table>
<thead>
<tr>
<th>Method</th>
<th>TC10</th>
<th>TC12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LBP_{riu2}$</td>
<td>97.19</td>
<td>86.67</td>
</tr>
<tr>
<td>$LBP_{riu2}$/VAR</td>
<td>98.33</td>
<td>87.01</td>
</tr>
<tr>
<td>LBP-HF</td>
<td>96.59</td>
<td>92.52</td>
</tr>
<tr>
<td>VZ_M8R8</td>
<td>87.75</td>
<td>87.49</td>
</tr>
<tr>
<td>VZ_Patch</td>
<td>94.11</td>
<td>92.64</td>
</tr>
<tr>
<td>proposed</td>
<td>99.18</td>
<td>96.20</td>
</tr>
</tbody>
</table>

Here, the VZ_Patch method used circular neighborhoods with radius 3 pixels [7](corresponding to a 7×7 patch) to overcome the rotation variant of the patch. Dominant orientation of each patch was computed to adjust their orientations. Table 1 presents the results of methods on TC10 and TC12 test suites. The highest scores 99.18% and 96.20% are obtained by the proposed method on TC10 and TC12, respectively. The results of our method are better than the results of $LBP_{riu2}$ which estimates the distribution of patterns in the whole image. These results validate that the $LBP_{riu2}$ distribution in small regions provide richer discriminant information than the $LBP_{riu2}$ distribution of the whole image. Fig. 2 gives an example of the misclassification with the $LBP_{riu2}$ method. The $LBP_{riu2}$ classified a texture from the 7th class into the 14th class because of their similar $LBP_{riu2}$ histograms. However, the distributions of $LBP_{riu2}$ histogram textons of two textures are very different. The bins 120~139 have large values in the texton histogram of the first texture (from the 7th class), while the bins 260~279 have large values in the texton histogram of the second texture(from the 14th class).

3.2. Experiment on CUReT database

The CUReT database has 61 materials, and 205 images of each material are acquired at different viewpoints and illumination. Fig. 3 shows a material under different conditions. There are 118 images that have been selected from a viewing angle. Fig. 2 gives an example of the misclassification with the $LBP_{riu2}$ method. The $LBP_{riu2}$ classified a texture from the 7th class into the 14th class because of their similar $LBP_{riu2}$ histograms. However, the distributions of $LBP_{riu2}$ histogram textons of two textures are very different. The bins 120~139 have large values in the texton histogram of the first texture (from the 7th class), while the bins 260~279 have large values in the texton histogram of the second texture(from the 14th class).

![Fig. 2](image_url)

**Fig. 2.** The two similar textures are misclassified by the $LBP_{riu2}$, but correctly classified by the proposed method.
angle less than 60°. In order to have the same experimental samples with literature[6, 7], among the 118 images from each texture class, we selected the 92 images large enough to be cropped to 200×200. Two test suites, FT23 and ET23, were used to test the proposed method. The training and testing samples are very different in FT23 test suite but similar in ET23 test suites.

1. FT23 test suite: The 92 images of each class were partitioned into two parts: the first 23 images of each class formed the training set, and other images grouped the testing set. Hence, there were 1403 (61×23) models and 4209(61×69) testing samples in total.

2. ET23 test suite: A training sample was selected every four samples among 92 images of each class. Thus, the 23 training sample of each class was evenly distributed among their images. Similarly, other images grouped to the testing set. And there were also 1403 (61×23) models and 4209(61×69) testing samples.

Fig. 3. The same material is captured under different illumination and viewpoints.

Table 2 gives the results of all the methods on FT23 and ET23 test suites. The proposed method gets the best score (78.88%) on FT23, while VZ_MR8 method obtains the highest score (94.95%) on ET23 (note that our score is still quite comparable). The VZ_MR8 and VZ_Patch perform well when the training samples are similar to testing samples, as on ET23, but on FT23 the performance of VZ_MR8 and VZ_Patch method deteriorate rapidly because of the dramatic change of illumination and viewpoint. In this situation, the LBP based methods have been less affected because of the robustness to illumination of the LBP operator. Therefore, the proposed method gets the highest score on FT23, even the original LBP<sub>pris2</sub> method that simply uses the global histogram of the entire image has a similar result as VZ_MR8 on FT23.

4. CONCLUSIONS

In this paper, we showed that the LBP distribution of different small regions in the same texture image might be different. The conventional LBP methods just computed the LBP histogram in the whole image and neglected the LBP statistic feature in local regions. In order to use full potential of local binary patterns, we extracted the rotation invariant LBP histograms in small regions to learn the LBP texton dictionary for texture description. All the texture images were represented with the occurrence of textons, not the occurrence of local binary patterns. The experimental results on public databases validated the advantage of proposed method.

5. REFERENCES