Rotation Invariant Texture Classification Using Binary Filter Response Pattern (BFRP)

Zhenhua Guo, Lei Zhang, and David Zhang

Biometrics Research Centre, Department of Computing, the Hong Kong
Polytechnic University, Hong Kong
{cszguo,cslzhang,csdzhang}@comp.polyu.edu.hk

Abstract. Using statistical textons for texture classification has shown great success recently. The maximal response 8 (MR8) method, which extracts an 8-dimensional feature set from 38 filters, is one of state-of-the-art rotation invariant texture classification methods. However, this method has two limitations. First, it require a training stage to build a texton library, thus the accuracy depends on the training samples; second, during classification, each 8-dimensional feature is assigned to a texton by searching for the nearest texton in the library, which is time consuming especially when the library size is big. In this paper, we propose a novel texton feature, namely Binary Filter Response Pattern (BFRP). It can well address the above two issues by encoding the filter response directly into binary representation. The experimental results on the CUReT database show that the proposed BFRP method achieves better classification result than MR8, especially when the training dataset is limited and less comprehensive.

Keywords: Texture Classification, Texton, MR8, LBP.

1 Introduction

Texture analysis is an active research topic in the fields of computer vision and pattern recognition. Generally speaking, it involves four basic problems: classifying images based on texture content; segmenting an image into regions of homogeneous texture; synthesizing textures for graphics applications; and establishing shape information from texture cue [1].

In the early stage, researchers devoted themselves into extracting statistical feature to classify texture images, such as the co-occurrence matrix method [2] and the filtering based methods [3]. These methods could achieve good classification results if the training and testing samples have similar or identical orientation. In real applications, rotation invariance is a critical issue to be solved. Kashyap and Khotanzad [4] were among the first researchers to study rotation-invariant texture classification by using a circular autoregressive model. Later, many other models were explored, including the multiresolution autoregressive model [5], hidden Markov model [6], and Gaussian Markov random field [7]. Jafari-Khouzani and Soltanian-Zadeh proposed to use Radon transform to estimate the texture orientation and extract wavelet energy features for texture classification [8]. In [11], Ojala et al. proposed to use the Local Binary

Pattern (LBP) histogram for rotation invariant texture classification. LBP is a simple but efficient operator to describe local image patterns. Recently, Varma and Zisserman [9] proposed a statistical learning based algorithm, namely Maximal Response 8 (MR8), using a group of filter banks, where a rotation invariant texton library is first built from a training set and then an unknown texture image is classified according to its texton distribution. Under the same framework, Varma and Zisserman [10] proposed a new statistical learning based algorithm. Instead of using filter bank, pixel gray value is used directly to represent local region. Scale and affine invariant texture classification is anther active research topic, and some pioneer work have been done by using fractal analysis [13-14] and affine adaptation [12].

As a state-of-the-art rotation invariant texture classification method, MR8 has achieved good accuracy on public database [9]. However, this method requires a training step to learn a feature dictionary. The learned dictionary thus depends on training samples. For example, it may suffer from generality when the training sample set is limited. Furthermore, to obtain a statistical histogram for a given image, it requires an additional step to assign the local region with a texton by searching the closest one from the trained library. This step is time consuming when the library size is big.

To solve the above problems, in this paper we propose a new feature extraction operator, namely Binary Filter Response Pattern (BFRP), which could be viewed as a binary version of MR8. After filtering the input image, BFRP converts the filter response into binary strings directly, instead of preserving the real values of filter response. Hence, each local region is assigned with a predefined texton. Such a binary presentation is not only fast to compute, but also can retain more discriminant information for classification.

The rest of the paper is organized as follows. Section 2 introduces the proposed BFRP and the dissimilarity metric. Section 3 reports the experimental results on a representative texture database. Section 4 gives the conclusion and future work.

2   Binary Filter Response Pattern

The MR8 filter bank used in [9] is employed in the proposed binary filter response pattern (BFRP) method. The MR8 filter bank consists of 38 filters, which are shown in Fig.1. To achieve rotation invariance, the filters are implemented at multiple
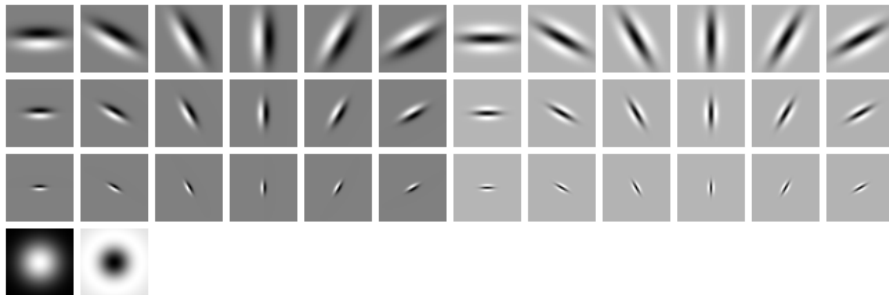

Fig. 1. The MR8 filter bank consists of a series of anisotropic filters (an edge and a bar filter at 6 orientations and 3 scales), and 2 rotationally symmetric ones (a Gaussian and a Laplacian of Gaussian) [9]

orientations and multiple scales. At each scale only the maximal response among the different orientations is kept. The final response at each position is an 8-dimension feature vector (3 scales for the edge and bar filters, plus 2 isotropic filters).

2.1 BFRP Feature Extraction

As shown in Fig.2, some local regions may have multiple dominant orientations. The magnitude of the filter response at each angle could be treated as a confidence measurement in the feature occurring at that orientation [9]. Thus we define the BFRP for multiple orientations as

$$BFRP(\text{x,y}) = \sum_{i=0}^{N-1} s(\, F_{\theta_i}(\text{x,y}) \,)2^i \;, \;\; s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}, \;\; \theta_i = \frac{i}{N}\pi \;, \; i=0,1,\ldots,\text{N-1} \qquad (1)$$

$$F_{\theta_i}(\text{x,y}) = f_{\theta_i} * I(\,x,y\,) \qquad (2)$$

where f_{θ_i} is the filter with orientation θ_i, $*$ is the convolution operation. N is the number of filters at one scale ($N=6$ for bar and edge filters, $N=1$ for Gaussian and Laplacian filters). I is the input image.
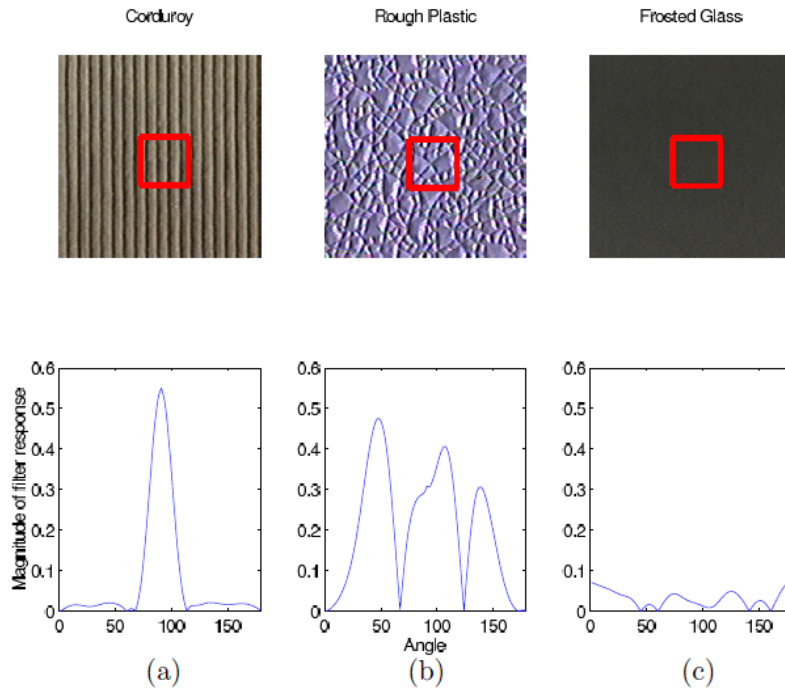

Fig. 2. The top row shows 3 texture images. The central image patch (highlighted by red rectangle) is matched with an edge filter at all orientations. The magnitude of the filter response versus the orientation is plotted in the bottom row [9].

To achieve rotation invariance, we shift the BFRP until the least value of the binary string is obtained. Similar to the definition of rotation invariant LBP code [11], we can define the rotation invariant BFRP and denote this binary code as $BFRP^{ri}$ (the superscript "*ri*" means the use of rotation invariant patterns). The rotation invariance of a binary code is achieved by shifting a number of binary bits, to find the least bit string value [11]. Thus, for the 6-bit binary code of bar and edge filters, there will be 14 kinds of rotation invariant strings (the 14 rotation invariant strings are 000000, 000001, 000011, 000101, 000111, 001001, 001011, 001101, 001111, 010101, 010111, 011011, 011111, and 111111). While for the 1-bit binary code of Gaussian and Laplacian filters, there will be 2 kinds of string only. With $BFRP^{ri}$, the filtering output at each position is a 8-dimensional vector, and there are 30,118,144 (14*14*14*14*14*14*2*2) kinds of patterns in total. Such a dimension is too large to build histogram and it will bring computation issue. To reduce the feature size, we empirically divide the 38 filters into 2 groups as show in Fig. 3. Thus for each image, only two 4-dimensional histograms need to be built and then the 2 histograms are concatenated. The final histogram size is reduced to 10,976 (14*14*14*2*2).
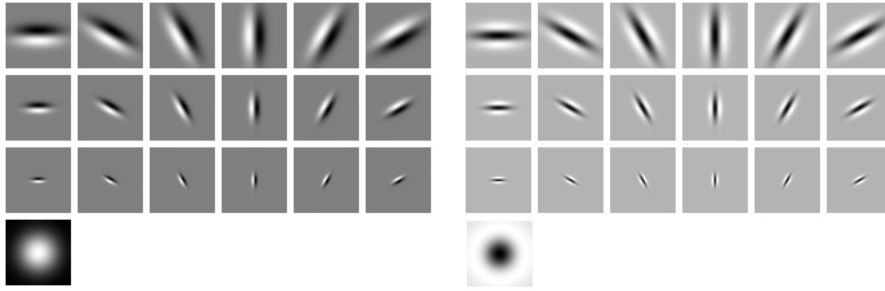

Fig. 3. Divide the MR8 filter banks into two equally groups

To further reduce the number of patterns at each scale, similar to the "uniform" LBP [11], the "uniform" BFRP is defined as:

$$BFRP^{riu2} = \begin{cases} \sum_{i=0}^{N-1} s(F_{\theta_i}) & if \ \ U(BFRP) \leq 2 \\ N+1 & \text{otherwise} \end{cases} \qquad (3)$$

$$U(BFRP) = \left| s(F_{\theta_{N-1}}) - s(F_{\theta_0}) \right| + \sum_{i=1}^{N-1} \left| s(F_{\theta_i}) - s(F_{\theta_{i-1}}) \right| \qquad (4)$$

The uniformity measure U is used to count the number of bit transition (bitwise 1/0 changes) in the bit string. By definition, exactly $N+1$ "uniform" binary patterns can occur in a circularly symmetric neighbor set of N binary bits, while the remaining (non-uniform) are grouped into a "miscellaneous" label (N+1).

The *BFRP*^{riu2} (superscript "*riu2*" means the use of rotation invariant "uniform" patterns that have *U* value of at most 2), has 8 distinct output values for 6-bit binary strings and 2 distinct output values for 1-bit binary strings. Using *BFRP*^{riu2}, the feature size is 2,048 (8*8*8*2*2) which is comparable with that in MR8 [9].

2.2 Dissimilarity Metric

The dissimilarity of sample and model histograms is a test of goodness-of-fit, which could be measured with a nonparametric statistic test. There are many metrics to evaluate the goodness between two histograms, such as histogram intersection, log-likelihood ratio, and chi-square statistic [11]. In this study, a test sample *T* is assigned to the class of model *L* that minimizes the chi-square distance:

$$D(T,L) = \sum_{n=1}^{N} \frac{(T_n - L_n)^2}{T_n + L_n} \tag{5}$$

where *N* is the number of bins, and T_n and L_n are the values of the sample and model image at the nth bin, respectively. In this paper, the nearest neighborhood classifier with chi-square distance is used to measure the dissimilarity between two histograms because it is equivalent to the optimal Bayesian classification [16] and get good performance for texture classification [17].

3 Experimental Results

In this section we will compare the proposed feature extraction scheme with MR8 and another non-training method (here, the "non-training" standing for the feature extraction is training free), LBP [11]. In MR8, 40 textons are clustered from each of the *c* texture classes using the training samples, and then a histogram based on the *c*40 textons is computed for each model and sample image. To more comprehensively evaluate the proposed method, in the experiments we list both the classification rates by using the *BFRP*^{ri} (rotation invariant BFRP) and *BFRP*^{riu2} ("uniform" rotation invariant BFRP) schemes.

The CURet database [18] is one of the largest yet most difficult databases for rotation invariant texture recognition. It contains 61 textures and each texture has 205 images obtained under different viewpoints and illumination directions. There are 118 images whose viewing angles (θ_v) are less than 60^0. Out of the 118 images, the 92 images, from which a sufficiently large region could be cropped (200*200) across all texture classes, are selected [9]. Before feature extraction, all the cropped regions are converted into grey scale and are normalized to have a mean of 0 and a standard deviation of 1 [9, 11]. Here, instead of computing error bar (i.e. mean and standard deviations of results calculated over multiple splits), we performed experiments on four different settings to simulated four situations:

1. T46A: The training set for each class was selected by taking one from every two adjacent images. Hence, there are 2,806 (61*46) models and 2,806 testing samples. Because in CUReT, adjacent images have similar appearance than far images, this setting is used to simulate the situation of large and comprehensive training set.

2. T23A: The training set for each class was selected by taking one from every four adjacent images. Hence, there are 1,403 (61*23) models and 4,209 (61*69) testing samples. This setting is used to simulate the situation of small but comprehensive training set.

3. T46F: The training set for each class was selected as the first 46 images. Hence, there are 2,806 models and 2,806 testing samples. This setting is used to simulate the situation of large but less comprehensive training set.

4. T23F: The training set for each class was selected as the first 23 images. Hence, there are 1,403 models and 4,209 testing samples. This setting is used to simulate the situation of small and less comprehensive training set.

The performance is evaluated by using the classification rate with chi-square distance and the nearest neighborhood classifier. The classification results by different operators with its associated feature size are listed in Table 1. The column for "T46A", "T23A", "T46F", "T23F" represent the four different test setup as described above.

From Table 1, we could make the following findings. First, under similar feature size, the proposed BFRP has better recognition performance than MR8. For example, $BFRP^{riu2}$ (2,048 dimension) achieves 0.39%, 0.33%, 1.14%, and 2.66% higher recognition rates than MR8 with its best settings (c=61 with feature size being 2,440 dimensions) for T46A, T23A, T46F, and T23F, respectively.

Second, because MR8 requires a training stage, which depends on the training samples, when the training sample is few and not comprehensive, the accuracy will drops quickly compared with the proposed feature extraction. For example, the accuracy decreases 18.32% (from T23A to T23F) for MR8 (c=61) while decreasing only 15.99% for $BFRP^{riu2}$. So the proposed method is more suitable for real applications where training samples are limited and not comprehensive.

Table 1. Classification rate (%) for the four settings using different schemes

Method	Feature Size	T46A	T23A	T46F	T23F
$LBP_{8,1}^{riu2} + LBP_{16,3}^{riu2} + LBP_{24,5}^{riu2}$ [15]	54	95.47	93.09	85.64	78.50
MR8 (c=61)	2440	97.65	96.15	88.70	77.83
MR8 (c=20)	800	96.79	94.80	86.89	75.98
MR8 (c=10)	400	95.55	93.87	84.50	74.82
$BFRP^{ri}$	10976	98.33	96.58	90.02	80.97
$BFRP^{riu2}$	2048	98.04	96.48	89.84	80.49
	Average	97.05	95.16	87.55	77.82

4 Conclusion

In the past decade, statistical texton has achieved great success for texture classification. However, it requires a training step and spends much cost on building the feature histogram. In this study, a training free rotation invariant feature extractor is proposed for texture classification. Using the same filter bank as in MR8, the proposed feature extractor requires much less computation cost on feature extraction but achieves better result than MR8. In the future we will investigate other statistical texton based operators, such as local fractal [14] and local patch [10], and try to extract their training free counterpart.

Acknowledgments

References

1. Tuceryan, M., Jain, A.K.: Texture analysis. In: Chen, C.H., Pau, L.F., Wang, P.S.P. (eds.) Handbook of pattern recognition and computer vision, ch. 2, pp. 235–276 (1993)
2. Haralik, R.M., Shanmugam, K., Dinstein, I.: Texture features for image classification. IEEE Trans. on Systems, Man, and Cybertics 3(6), 610–621 (1973)
3. Randen, T., Husy, J.H.: Filtering for texture classification: a comparative study. IEEE Trans. PAMI 21(4), 291–310 (1999)
4. Kashyap, R.L., Khotanzed, A.: A model-based method for rotation invariant texture classification. IEEE Trans. on PAMI 8(4), 472–481 (1986)
5. Mao, J., Jain, A.K.: Texture classification and segmentation using multiresolution simultaneous autoregressive models. Pattern Recognition 25(2), 173–188 (1992)
6. Wu, W.R., Wei, S.C.: Rotation and gray-scale transform-invariant texture classification using spiral resampling, subband decomposition, and hidden Markov model. IEEE Trans. IP 5(10), 1423–1434 (1996)
7. Deng, H., Clausi, D.A.: Gaussian MRF rotation-invariant features for image classification. IEEE Trans. on Pattern Analysis and Machine Intelligence 26(7), 951–955 (2004)
8. Jafari-Khouzani, K., Soltanian-Zadeh, H.: Radon transform orientation estimation for rotation invariant texture analysis. IEEE Trans.PAMI 27(6), 1004–1008 (2005)
9. Varma, M., Zisserman, A.: A statistical approach to texture classification from single images. International Journal of Computer Vision 62(1-2), 61–81 (2005)
10. Varma, M., Zisserman, A.: A statistical approach to material classification using image patch exemplars. IEEE Trans. PAMI (to appear)
11. Ojala, T., Pietikäinen, M., Mäenpää, T.T.: Multiresolution gray-scale and rotation invariant texture classification with Local Binary Pattern. IEEE Trans. PAMI 24(7), 971–987 (2002)
12. Lazebnik, S., Schmid, C., Ponce, J.: A sparse texture representation using local affine regions. IEEE Trans. PAMI 27(8), 1265–1278 (2005)

13. Xu, Y., Ji, H., Fermuller, C.: A projective invariant for texture. In: International Conference on Computer Vision and Pattern Recognition, pp. 1932–1939 (2005)
14. Varma, M., Garg, R.: Locally invariant fractal features for statistical texture classification. In: International Conference on Computer Vision (2007)
15. Pietikäinen, M., Nurmela, T., Mäenpää, T., Turtinen, M.: View-based recognition of real-world textures. Pattern Recognition 37(2), 313–323 (2004)
16. Varma, M., Zisserman, A.: Unifying statistical texture classification framework. Image and Vision Computing 22(14), 1175–1183 (2004)
17. Puzicha, J., Buhmann, J.M., Rubner, Y., Tomasi, C.: Empircal evaluation of dissimilarity measures for color and texture. In: International Conference on Computer Vision, pp. 1165–1172 (1999)
18. Dana, K.J., van Ginneken, B., Nayar, S.K., Koenderink, J.J.: Reflectance and texture of real world surfaces. ACM Trans. on Graphics 18(1), 1–34 (1999)