Yongsheng Dong and Jinwen Ma*

Abstract.

L_1

K

Keywords:

1

Texture classification plays an important role in computer vision with a wide variety of applications. Examples include classification of regions in satellite images, automated inspection, medical image analysis and document image processing. During the last three decades, numerous methods have been proposed for image texture classification and retrieval [1]-[13]. Among these approaches, wavelet-based methods may be the most popular due to the multiresolution and orientation representation of wavelets which is consistent with the human visual system [10].

However, two-dimensional wavelets are only good at catching point discontinuities but do not capture the geometrical smoothness of the contours. As a newly developed two-dimensional extension of the wavelet transform using multiscale and directional filter banks, the contourlet transform can effectively capture the intrinsic geometrical structure that is key in visual information, because the contourlet expansion can achieve the optimal approximation rate for piecewise smooth functions with C^2 contours in some sense [14]. Recently, the contourlet transform has been successf

Of course, the contourlet transform also provides us an efficient tool to extract the features from a texture image for texture classification.

Modeling wavelet detail subband coefficients via the Product Bernoulli Distributions (PBD) [11]-[12] has received a lot of interest. The PBD model makes use of a binary bit representation for wavelet subband histograms and the so-called Bit-plane Probability (BP) signature is constructed based on the model parameters. Essentially, the main merits of BP approach are its efficiency for signature extraction and similarity measurement based on Euclidean metric, and the statistical justification of the model parameters for use in image processing applications [10]-[11]. However, it has two main disadvantages. First, the wavelet transform used in [11] cannot capture directional information and then the wavelet coefficients don't represent a texture image well. So the recognition performance is not satisfying. Second, the minimum distance classifier used in [11] doesn't work well because the BP signature is obtained by concatenating all the bit-plane probabilities of all high-pass subbands, and the distance between a new image and a texture class is obtained by the $weighted - L_1$ distance of the BP signature of the test sample and the mean of the BP signatures of all training samples in each class.

Motivated by the advantages and disadvantages of PBD, we propose a new method for texture classification using contourlet transform and PBD together. More specifically, this paper makes the following contributions. First, we use product Bernoulli distributions to model the contourlet coefficients instead of wavelet coefficients. Second, we present a new distance of two images, which is measured by summing up all the $weighted - L_1$ metrics of the bit-plane probabilities of the corresponding subbands. Finally, we apply the PBD model in the contourlet domain to supervised texture classification through the K-nearest neighbor classifier, and experimental results on large texture datasets reveal that our proposed method with the use of the new distance performs better than the

Due to its cascade structure accomplished by combining the Laplacian pyramid (LP) with a directional filter bank (DFB) at each scale, multiscale and directional decomposition stages in the contourlet transform are independent of each other. Therefore, one can decompose each scale into any arbitrary power of two's number of directions, and different scales can be decomposed into different numbers of directions. Therefore, it can represent smooth edges in the manner of being close to the optimal efficiency. Fig. 1 shows an example of the contourlet transform on the "Barbara" image. For the visual clarity, only two-scale decompositions are shown. The image is decomposed into two pyramidal levels, which are then decomposed into four and eight directional subbands, respectively.

More recent developments and applications on the contourlet transform can be found in [15],[16] and [20].

Fig. 1.

3 P

3.1 Product Bernoulli Distributions

For L-scale contourlet decompositions of a given texture image, we find the average amplitude of the coefficients increases almost exponentially with the scale i ($i = 1, 2 \cdots, L$). Hence, to model the contourlet coefficients at different scales uniformly, we regularize them by multiplying the factor $1/2^i$ to those in the high-pass directional subbands at the i-th scale, and multiplying the factor $1/2^{2L}$ to those in the low-pass subband. For simplicity, the contourlet coefficients in the following will represent the regularized coefficients without explanation.

Considering one particular contourlet subband, we quantize each coefficient into n bits using deadzone quantization that step size equals 1.0 [11]. It follows that each of the quantized contourlet coefficients can be expanded into n binary bit-planes. Thus, we express it as a random variable:

$$Y = \sum_{i=0}^{n-1} 2^i Y_i, \tag{1}$$

where Y_i is a random variable representing the i-th binary bit of the coefficient. That is, each bit-plane is composed of either 0 or 1, and the joint probability distribution of the quantized coefficients is $P(Y = y) = P(Y_0 = y_0, Y_1 = y_1, \cdots, Y_{n-1} = y_{n-1})$ where $y_i \in \{0, 1\}$ is the i-th binary bit of y. If we assume Y_i's are statistically independent variables and denote the model parameter by $p_i = P(Y_i = 1)$, the joint distribution can be written as a product of Bernoulli distributions (PBD) [11]:

$$f^{PBD}(Y = y) = \prod_{i=0}^{n-1} p_i^{y_i}(1 - p_i)^{1-y_i}, \tag{2}$$

which can be characterized by the bit-plane probabilities: $P = (p_0, p_1, \cdots, p_{n-1})$.

In this way, for a given particular contourlet subband with a set of absolute quantized coefficients $\boldsymbol{y} = (y^1, y^2, \cdots, y^k)$ where $y^j \in Z^+$ is the j-th component of $\boldsymbol{y}$ and Z^+ denotes the set of all the nonnegative integer numbers, the likelihood function of $\boldsymbol{y}$ can be defined as

$$\tilde{L}(\boldsymbol{y}; P) = \log \prod_{j=1}^{k} f^{PBD}(Y = y^j; P) = \log \prod_{j=1}^{k} \prod_{i=0}^{n-1} p_i^{y_i^j}(1 - p_i)^{1-y_i^j}, \tag{3}$$

where y_i^j is the i-th binary bit of the j-th component of $\boldsymbol{y}$ and

$$P = (p_0, p_1, \cdots, p_{n-1}). \tag{4}$$

Thus, the ML estimator [11] of P can be obtained by

$$\frac{\partial \tilde{L}(\boldsymbol{y}; P)}{\partial p_i} = 0, \tag{5}$$

namely, $\hat{p}_i = \frac{1}{k}\sum_{j=1}^{k} y_i^j$, where $i = 0, 1, \cdots, n-1$. That is, the ML estimator of the model parameter is equivalent to the probabilities of one-bit occurrence for each of the bit-planes. Therefore, we can compute the bit-plane probabilities (BP) for each contourlet subband using the above ML estimators. As we all know, a sufficient statistic for a model parameter is a statistic that captures all possible information in the data about the model parameter. In the same manner as in [11], the sufficiency of the parameter estimators can also be proved by the Fisher-Neyman factorization theorem [11].

3.2 Discrepancy Measurement and *K*-Nearest Neighbor Classifier

Once the bit-plane probabilities (BP) of all subbands are obtained for every texture, we can compare the corresponding BPs of two subbands using a metric. In [11], it has been investigated and demonstrated that the $Relative - L_1$ (RL1) distance is suitable for comparing BPs. Hence, we still use RL1 as the metric of two BPs P^1 and P^2, which is given by

$$RL_1(P^1, P^2) = \sum_{i=0}^{n-1} \frac{|p_i^1 - p_i^2|}{1 + p_i^1 + p_i^2} \tag{6}$$

where $P^1 = (p_0^1, p_1^1, \cdots, p_{n-1}^1)$ and $P^2 = (p_0^2, p_1^2, \cdots, p_{n-1}^2)$. Note that the RL1 distance is a weighted L_1 one.

For two given images I_1 and I_2, we can obtain M contourlet subbands $(B_1^{I_1}, B_2^{I_1}, \cdots, B_M^{I_1})$ and $(B_1^{I_2}, B_2^{I_2}, \cdots, B_M^{I_2})$, respectively, after having implemented an L-level contourlet transform on them, and then define the distance between the two images by

$$DL(I_1, I_2) = \sum_{j=1}^{M} d_j, \tag{7}$$

where $d_j = RL_1(P_j^1, P_j^2)$ is the $Relative - L_1$ distance between the two BPs P_j^1 and P_j^2 corresponding to the subbands $B_j^{I_1}$ and $B_j^{I_2}$, respectively for $j = 1, 2, \cdots, M$.

Given a single test sample I^* and a training set, we will utilize the K-nearest-neighbor classifier to perform texture classification. In particular, we compare I^* with each training sample, and then assign it to the class to which the majority of these k nearest neighbors belong. This classifier performs better than the minimum distance classifier used in [11], which will be demonstrated in the following section.

4 a

In this section, various experiments are carried out to demonstrate our proposed method for texture classification. In our experiments, we select the pyramid and directional filters by the "9-7" filters in the contourlet transform, which are the biorthogonal wavelet filters. In addition, we impose that each image is decomposed into one low-pass subband and four high-pass subbands at four pyramidal levels. The four high-pass subbands are then decomposed into four, four, eight, and eight directional subbands, respectively. It follows that the total number of directional subbands, M, is 25. For texture images, $n = 8$ bits are sufficient for subband coefficients. For the sake of clarity, we refer to our proposed method based on the bit-plane probability model in the contourlet domain and K-NN classifier as BPC+KNN.

4.1 Performance Evaluation

We first evaluate our method for texture classification on a typical set of 64 grey 640×640 images (shown in Fig. 2 and denoted by Set-1) from the Brodatz database [21], which was also used in [23].

In the experiments on Set-1, each image is divided into 25 128×128 nonoverlapping patches, and thus there are totally 1600 samples available. We select N_{tr} training samples from each of 64 classes and let the other samples for test with $N_{tr} = 2, 3, \cdots, 24$. The partitions are furthermore obtained randomly and the average classification accuracy rate (ACAR) is computed over the experimental results on 10 random splits of the training and test sets at each value of N_{tr}.

Fig. 2. The set of 64 texture images used in [23]

For Set-1, we compare our proposed BPC+KNN with two other methods. The “rst is the method based on the bit-plane probability model in the wavelet domain and minimum distance classi“er (called BPW+MD)[11]. By this approach, the distance between a test sample and a given class is de“ned by the L_1 distance between the input BP signature of the test sample and the mean of the BP signatures of all training samples in the class. Once the distances between the test sample and each class are obtained, the label of the class that has the minimum distance from the test sample is assigned to the test sample. The second method is BPC+MD, which is the same as BPC+KNN but the minimum distance (MD) classi“er. For the MD classi“er, the distance between a test sample and a given class is de“ned as the mean of the ones, de“ned by DL, between the test sample and all training samples in the texture class.

Fig. 3 plots the average classi“cation accuracy rates (with error bars) of BPC+KNN, BPC+MD, and BPW+MD with respect to the number of training samples N_{tr}. As can be seen, the ACAR of BPC+KNN increases monotonically with the number of training samples. However, the ACAR of BPW+MD does not have the same regularity as that of BPC+KNN. We can also see that BPC+MD performs better than BPW+MD by about 2.0%-4.0% for each value of N_{tr}, which implies PBDs in the contourlet domain outperforms those in the wavelet domain. BPC+KNN ($K = 1$) slightly outperforms BPC+KNN (K=3) and performs better than BPC+MD by 1.9%-4.5% for each value of N_{tr}, which implies that the KNN classi“er outperforms the MD classi“er for the PBD model. Note that the errors are also shown in Fig. 3 where each error bar is a distance of one standard deviation above and below the average classi“cation accuracy rate. All the values of standard deviation of BPC+KNN with $K = 1$ and $K = 3$ at each value of N_{tr} are about 0.50%, which are slightly less than the average value of standard deviation of BP Method, 1.54%. In other words, the variation of

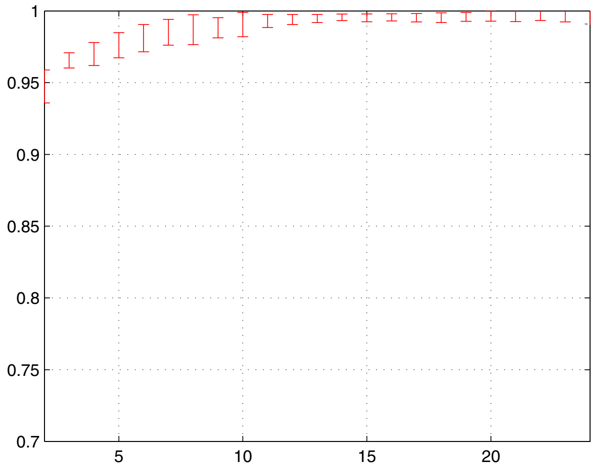
1
0.95
0.9
0.85
0.8
0.75
0.7
5
10
15
20

Table 1.

	K		
	. ± .	. ± .	
	. ± .	. ± .	.

Table 2.

	K	K		
	.	.	.	.

method (called MR Method) [17] on the Vistex dataset [22] of 129 512×512 texture images (denoted by Set-3), which has also been used in [17]. By the MR Method, the texture features are extracted by the M-band ridgelet transform, which is obtained by combining the M-band wavelet with the ridgelet. Then, the support vector machine classifier is used to perform supervised texture classification.

Each texture image is divided into 16 128×128 non-overlapping patches. We select 8 training samples and compute the ACAR over the experimental results on 10 random splits of the training and test sets. The average classification results of these methods are listed in Table 1. It can be seen from Table 1 that BPC+KNN ($K = 1$) outperforms MR Method and BPW+MD by 1.42% and 14.24%, respectively, on the large dataset.

4.2 Computational Cost

We further compare our proposed method with the other methods on computational cost. All the experiments conducted here have been implemented on a workstation with Intel(R) Core(TM) i5 CPU (3.2GHz) and 3G RAM in Matlab environment.

Tabel 2 reports the time for texture classification (TTC) and ACARs using the BPC + KNN ($K = 1$), BPC + KNN ($K = 3$), BPC + MD and BPW + MD approaches on the 64 texture dataset. The number of training samples used in the experiments is 8. For this dataset, the BPC + KNN ($K = 3$) method is the most efficient. In contrast, BPW + MD is the most time-consuming method among them. The TTC using BPC + KNN ($K = 1$) is 206.17 s, which is about 3 times faster than the BPW + MD. In addition, BPC + KNN ($K = 1$) is also slightly more efficient than BPC + MD. If we take into account the TTC and ACAR, the results clearly show that BPC + KNN ($K = 1$) outperforms the other methods.

5

We have investigated the distribution of the coefficients in each contourlet subband and tried to use the product Bernoulli distributions for modeling them. We then apply the PBD model with the use of KNN classifier to supervised texture classification. The various experiments have shown that our proposed method considerably improves the texture classification accuracy in comparison with the current state-of-the-art method based on product Bernoulli distributions in the wavelet domain as well as the method based on the M-band ridgelet transform.

This work was supported by the Natural Science Foundation of China for grant 60771061.

http://www.ux.uis.no/~tranden/brodatz.html

http://ww.vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html