Wavelet-Based Image Texture Classification Using Local Energy Histograms

Yongsheng Dong and Jinwen Ma

***Abstract*—In this letter, we propose an efficient one-nearest-neighbor classifier of texture via the contrast of local energy histograms of all the wavelet subbands between an input texture patch and each sample texture patch in a given training set. In particular, the contrast is realized with a discrepancy measure which is just a sum of symmetrized Kullback–Leibler divergences between the input and sample local energy histograms on all the wavelet subbands. It is demonstrated by various experiments that our proposed method obtains a satisfactory texture classification accuracy in comparison with several current state-of-the-art texture classification approaches.**

***Index Terms*—Energy histogram, one-nearest-neighbor classifier, symmetrized Kullback–Leibler divergence (SKLD), texture classification, wavelet transform.**

I. Introduction

TEXTURE classification plays an important role in computer vision with a wide variety of applications. During the last three decades, numerous methods have been proposed for image texture classification or retrieval. These methods can be broadly divided into four categories [1], namely structural methods, statistical methods, model-based methods and filter-based methods. As a special subcategory of filter-based methods [1]–[13], wavelet-based multiresolution methods have recently received extensive interest in the literature, which can be considered to follow two major streams: model-based methods [3]–[7] and feature-based methods [1], [8]–[12].

The distribution characteristic of wavelet subband coefficients has played an important role on texture classification in recent years [1]–[7]. The used models include the Characteristic Generalized Gaussian Density (CGGD) model [3], the Bit-plane Probability (BP) model [5], [6], the Refined Histogram (RH) [4], the Generalized Gamma Density (GΓD) model [7], and so on. The wavelet coefficients are generally modeled under the independence assumption that disregards the intercoefficient dependencies. Thus, the valuable local information that those dependent coefficients in a given neighborhood own together will be neglected.

Manuscript received December 06, 2010; revised January 20, 2011; accepted January 26, 2011. Date of publication February 04, 2011; date of current version February 17, 2011. This work was supported by the Natural Science Foundation of China under Grant 60771061. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. H. Vicky Zhao.

The authors are with the Department of Information Science, School of Mathematical Sciences and LMAM, Peking University, Beijing 100871, China (e-mail: jwma@math.pku.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2011.2111369

In the feature-based methods, the total energy of each high-pass wavelet subband is a commonly used statistical feature for texture classification [11], [12]. Moreover, the local energy features in each high-pass subband is also extracted and used to perform texture classification [1]. However, all the above wavelet-based methods ignore the low-pass subband in performing texture classification.

To utilize the local energy feature distributions of high-pass and low-pass wavelet subbands for texture classification, we here propose a special local energy histogram (LEH) to describe the distribution of the coefficients in each wavelet subband. In this way, we define a new discrepancy measure between two images by summing up all the symmetrized Kullback–Leibler divergences between two LEHs on the corresponding subbands. According to this discrepancy measure, a one-nearest-neighbor classifier is then built for supervised texture classification. It is demonstrated by the experiments that our proposed method has a satisfactory classification performance in comparison with several current state-of-the-art texture classification approaches.

The rest of this letter is organized as follows. Section II presents our proposed texture classification method. Its experimental results and comparisons are given in Section III. Finally, we conclude briefly in Section IV.

II. Proposed Texture Classification Method

A. Local Energy Feature Extraction in Wavelet Domain

For an L-level wavelet decomposition, we obtain $3L$ high-pass subbands $(B_1, B_2, \ldots, B_{3L})$ and one low-pass subband B_{3L+1}. Then we extract the local (Norm-1) energy features on $S \times S$ neighborhoods in each subband. Typically, in the j-th high-pass subband of size $\Omega_i^j \times \Omega_i^j$ at the i-th scale, the local energy features can be defined by

$$E_{\mathrm{Loc}}^{i,j}(l,k) = \frac{1}{S^2} \sum_{u=1}^{S} \sum_{v=1}^{S} |w_{i,j}(l+u-1, k+v-1)| \quad (1)$$

where $1 \le l,k \le \Omega_i^j - S + 1$ and $w_{i,j}(m,n)$ is the wavelet coefficient at location (m,n) in the subband. The local energy features in the low-pass subband, denoted by E_{Loc}^{L} for clarity, are also extracted in the same manner according to (1).

Note that all the above local energy features are non-negative. On the other hand, the average amplitude of the local energy values increases almost exponentially with the scale i. To make a uniform measure for those local energy features at different scales, we regularize $E_{\mathrm{Loc}}^{i,j}$ by multiplying the factor $1/2^i$ to them. For the low-pass subband, we multiply the factor $1/4^L$ to E_{Loc}^{L} due to that the average amplitude of the local energy feature values in the low-pass subband is much higher than that

in each high-pass subband. For simplicity, the local energy features in the following will be considered as the regularized local energy features without explanation.

B. Local Energy Histogram (LEH)

1) Definition: Given a particular wavelet subband with M local energy features $E = (e_1, e_2, \ldots, e_M)$, there certainly exists an positive integer N_0 such that the maximum $e_{\max} < 2^{N_0}$. Let $\Delta_n = [2^a(n-1), 2^a n)$, $n = 1, 2, \ldots, N$, where $N = 2^{N_0 - a}$ and $0 \le a \le N_0$, which is called the bin-width index. Note that $[0, 2^{N_0}) = \bigcup_{n=1}^{N} \Delta_n$. It follows that, given any e_m, there must exist an positive integer n_0 $(1 \le n_0 \le N)$ such that $e_m \in \Delta_{n_0}$. Then we define a local energy histogram (LEH) as a discrete function $p(\Delta_n) = p_n = m_n/M$, where m_n is the number of the local energy features appearing in Δ_n. Note that the defined local energy histogram is a normalized one, which can be used to model the probability density function of the local energy features. Hence, we characterize E by $P = (p_1, p_2, \ldots, p_N)$, which is referred as the LEH signature.

2) Discrepancy Measure: Once the LEH signature $P = (p_1, p_2, \ldots, p_N)$ in each wavelet subband is obtained for every texture image in a given dataset, classification requires the comparison of LEHs for similarity. Without the loss of generality, we assume that all the LEH signatures have the same length. Because we can use a big enough number N_0 such that for any texture image in the considered dataset, the maximum local energy value in all the subbands is strictly less than 2^{N_0}. Note that we use the same bin-width index a fixed previously for all subbands in our experiments. To avoid $p_n = 0$ for some n and then prevent the divide by zero problems in the following discrepancy measure, we subtract η from p_{n^*}, where $n^* = \arg\max_n p_n$ and η is a sufficiently small positive real number such as $\eta = 10^{-16}$. That is, we let $p'_n = p_n + \eta/(N-1)$ if $n \ne n^*$, but $p'_{n^*} = p_{n^*} - \eta$. It follows that $\sum_{n=1}^{N} p'_n = \sum_{n=1}^{N} p_n = 1$. For the sake of clarity, the LEH signature $P = (p_1, p_2, \ldots, p_N)$ in the following will represent $P' = (p'_1, p'_2, \ldots, p'_N)$ without explanation. A well-known discrepancy measure between two probability distributions is the symmetrized Kullback–Leibler divergence (SKLD) [4]. So, we define the SKLD between two LEHs H and Q by

$$\mathrm{SKLD}(H, Q) = \sum_{n=1}^{N} p_n \log\left(\frac{p_n}{q_n}\right) + \sum_{n=1}^{N} q_n \log\left(\frac{q_n}{p_n}\right) \quad (2)$$

where p_n and q_n are the LEH signatures of H and Q, respectively.

Given two images I_1 and I_2, we can obtain $3L+1$ wavelet subbands $(B_1^{I_1}, B_2^{I_1}, \ldots, B_{3L+1}^{I_1})$ and $(B_1^{I_2}, B_2^{I_2}, \ldots, B_{3L+1}^{I_2})$, respectively, after having implemented an L-level wavelet transform on them, and then define the discrepancy measure between the two images by

$$TD = \sum_{i=1}^{3L+1} d_i = HD + d_{3L+1} \quad (3)$$

where $HD = \sum_{i=1}^{3L} d_i$, and $d_i = \mathrm{SKLD}(H_i^{I_1}, Q_i^{I_2})$ is the SKLD between the two LEHs $H_i^{I_1}$ and $Q_i^{I_2}$ corresponding to the subbands $B_i^{I_1}$ and $B_i^{I_2}$, respectively for $i = 1, 2, \ldots, 3L+1$. As

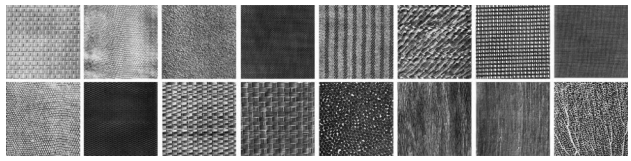

Fig. 1. Sixteen of the 80 Brodatz texture images.

d_i essentially measures the discrepancy between the subbands $B_i^{I_1}$ and $B_i^{I_2}$, we refer to it as a subband distance. It is clear that TD is the summation of all the subband distances. Note that we can also utilize the summation of all the high-pass subband distances, i.e., HD, to define the discrepancy measure between two images, which is equivalent to performing texture classification using only all the high-pass subbands. In the next section, we will demonstrate the importance of the low-pass subband to classification performance by comparing TD with HD.

C. One-Nearest-Neighbor Classifier

With the above discrepancy measure TD, we will utilize the one-nearest-neighbor classifier as our default classifier for its good performance. It follows that, given a training sample set and a test sample, we will compare the input test sample with all the training samples via the discrepancy measure and assign the class label of the closest training sample to it. In fact, it has been found by the experiments that, with the discrepancy measure TD, the one-nearest-neighbor classifier performs considerably better than the minimum distance classifier and the k-nearest-neighbor classifier with $k > 1$ for supervised texture classification.

III. Experimental Results

In this section, various experiments are carried out to demonstrate our proposed texture classification method, being compared with several current state-of-the-art texture classification approaches under different image texture environments.

In the experiments, the 3-level wavelet transform with the Daubechies 1 (db1) filter bank is used to decompose each sample patch. We also considered the range of the level from 1 to 4 in our experiments. The results showed that 3-level wavelet transform gives the best results. As for the selection of the parameter N_0, the experiments show that $N_0 = 10$ is big enough, that is, for any texture image in our datasets, the maximum local energy value in all the subbands is strictly less than 2^{10}. Therefore, we shall use $N_0 = 10$ in all the following experiments. As a matter of fact, the results will be almost the same if we select a larger integer number N_0.

A. Classification Performance

We first evaluate our texture classification method on a typical set of 80 grey 640×640 images (denoted by Set-1, and 16 of the 80 images are shown in Fig. 1) from the Brodatz texture database [14], which was also used in [17].

In the experiments on Set-1, each image is divided into 16 160×160 nonoverlapping patches, and thus there are totally 1280 samples available. We select N_{tr} training samples from each of 80 classes and let the other samples for test with $N_{tr} = 2, 3, \ldots, 8$. The partitions are furthermore obtained randomly

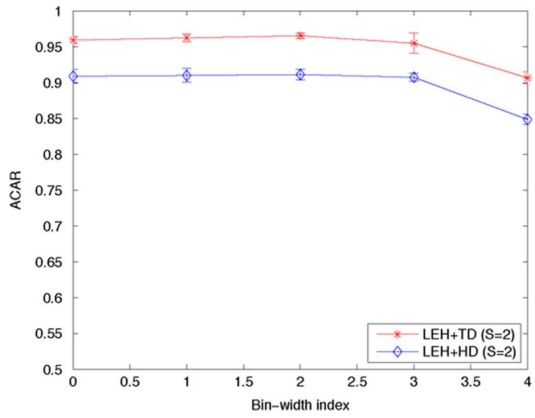

Fig. 2. Sketches of the average classification accuracy rates of LEH+TD ($S = 2$) and LEH+HD ($S = 2$) with respect to the bin-width index when the number of training samples is 8.

and the average classification accuracy rate (ACAR) is computed over the experimental results on ten random splits of the training and test sets for each value of N_{tr}.

For simplicity, we refer to our proposed method using the discrepancy measure TD as LEH+TD, and to the one-nearest-neighbor classifier based on the LEH and the discrepancy measure HD [see (3)] as LEH+HD. Fig. 2 plots the ACARs of LEH+TD and LEH+HD with respect to the bin-width index $(a = 0, 1, \ldots, 4)$ using the neighborhood size 2×2 (i.e., $S = 2$) for the local energy features in the case of $N_{tr} = 8$. Note that the error bars are also shown in Fig. 2 and the following figures where each error bar is a distance of one standard deviation above or below the average classification accuracy rate. From Fig. 2, we can see that the ACARs of LEH+TD and LEH+HD decrease with the bin-width index, a, and LEH+TD outperforms LEH+HD by 5.00%–6.00%, which implies the discrepancy measure TD performs much better than HD in each case of the bin-width index. In these two approaches, we can also see that the difference between the ACARs of $a = 0$ and $a = 4$ cases is 5.00%–6.50%, which implies the bin-width index a plays an important role to classification performance. Note that as a decreases from 3 to 0, the improvement on ACAR is insignificant. The main reason is that, as a becomes smaller, our LEHs can capture more details about the distribution of the local energy features, and thus, the ACAR increases. However, we only have more details on the peaks of the distribution when it is too small (e.g., $a = 0$), which is actually unnecessary. In such a case, the recognition performance cannot be significantly improved. Therefore, in the following experiments, we shall only use the bin-width index $a = 1$ to demonstrate our method.

We then investigate the sensitivity of the low-pass subband to classification performance when the neighborhood sizes are 2×2, 3×3 and 4×4 (that is, $S = 2$, 3, and 4). Figs. 3(a), (b) and (c) plot the ACARs (with error bars) of LEH+TD and LEH+HD with respect to the number of training samples N_{tr} for the three neighborhood sizes, respectively. As can be seen, the ACARs of LEH+TD and LEH+HD increase monotonically with the number of training samples no matter what the neighborhood size is. We can also see that LEH+TD outperforms LEH+HD by 4.00%–6.00% at each value of N_{tr} in each case of the neighborhood size, which implies the discrepancy measure

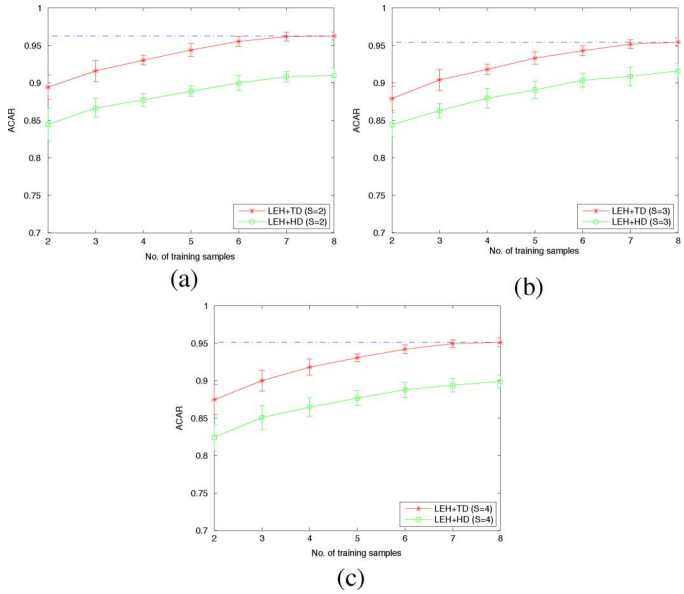

Fig. 3. Classification performances of LEH+TD and LEH+HD with respect to the number of training samples using the three neighborhood sizes: (a) $S = 2$; (b) $S = 3$; (c) $S = 4$.

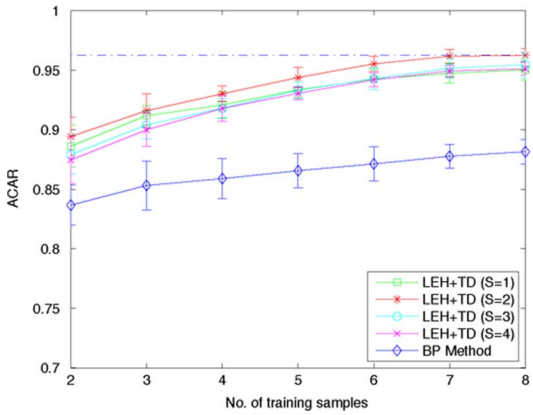

Fig. 4. Comparative classification performance with respect to the number of training samples.

TD performs better than HD. Therefore, the low-pass subband of one texture image plays an essential role in distinguishing different texture classes.

We finally compare our proposed texture classification method, LEH+TD, with the bit-plane probability (BP) signature based method [5], which is referred as BP Method. Fig. 4 shows the ACAR sketches of the two methods with respect to the number of training samples, N_{tr}. It can be clearly seen from Fig. 4 that the ACARs of LEH+TD with the three neighborhood sizes and BP Method increase monotonically with the number of training samples. We also see that LEH+TD ($S = 2$) performs slightly better than LEH+TD ($S = 1$), LEH+TD ($S = 3$) and LEH+TD ($S = 4$) by about 1.00%, and outperforms BP Method by about 6.00%–8.00% for each value of N_{tr}, which implies that the optimal selection of the neighborhood size is $S = 2$. Moveover, we compare LEH+TD ($S = 2$) with the Gaussian mixture model based method with the basis system [17] (known as GMM Method) in the case of $N_{tr} = 8$. The ACARs of GMM Method and LEH+TD ($S = 2$) are 93.44% and 96.25%, respectively. That

TABLE I
THE ACARS (%) AND TTC (IN SECONDS) OF THE THREE METHODS

	LEH+TD ($S = 2$)	CLBP_S	CLBP_S/M/C
ACAR	95.29	92.46	98.08
TTC	140.99	138.36	144.37

is, LEH+TD ($S = 2$) outperforms GMM Method by 2.81%. All the values of standard deviation of LEH+TD with the three neighborhood sizes at each value of N_{tr} are about 0.90%, which are slightly less than the average value of standard deviation of BP Method, 1.48%. In other words, the variation of the classification accuracy rates of LEH+TD for different number of training samples is small, which affirms the robustness of our proposed LEH+TD.

B. Comparisons With the Other Existing Methods

In this subsection, we firstly test LEH+TD ($S = 2$) on the Brodatz dataset [14] of 111 640×640 texture images (denoted by Set-2), and compare with BP Method and the independent component analysis (ICA) signature based method (known as ICA Method) [16]. Each texture image is divided into 25 128×128 non-overlapping patches. Then, we select 10 training patches from each class and put the remaining patches into the test set. The experiment is repeated over ten random splits of the training and test sets. The ACARs of LEH+TD ($S = 2$), BP Method, and ICA Method are 85.80%, 74.90%, and 80.7%, respectively. We can see that LEH+TD ($S = 2$) clearly outperforms the state-of-the-art BP Method and ICA Method on the entire Brodatz dataset.

To provide further justification of our proposed LEH+TD ($S = 2$), we also compare LEH+TD ($S = 2$) with two current methods based on the completed local binary pattern (CLBP), denoted by CLBP_S and CLBP_S/M/C as in [18], on the Vistex dataset [15] of 30 512×512 texture images. We leave more details and developments about the local binary pattern in [18] and [19]. Our experimental setting is the same as that used in [3]. The experiment is repeated over ten random splits of the training and test sets. All the experiments are implemented on a workstation with Intel(R) Core(TM) i5 CPU (3.2 GHz) in Matlab environment. Table I reports the ACARs and time for texture classification (TTC) of these methods. As can be seen, our proposed LEH+TD ($S = 2$) outperforms CLBP_S by 2.83% though CLBP_S/M/C performs better than LEH+TD ($S = 2$) by 2.79%. The ACARs of the three methods are all higher than that of the characteristic generalized Gaussian density based method [3], i.e., 88.1%. As far as the TTC is concerned, LEH+TD ($S = 2$) is slightly faster than CLBP_S/M/C. Therefore, our proposed LEH+TD ($S = 2$) is very competitive to the current state-of-the-art CLBP_S and CLBP_S/M/C approaches.

IV. CONCLUSION

We have investigated the supervised texture classification problem by contrasting the local energy histograms of all the wavelet subbands between an input texture patch and each sample texture patch in a given training set. The contrast is conducted with a discrepancy measure defined as a sum of the symmetrized Kullback–Leibler divergences between the input and sample local energy histograms on all the wavelet subbands, and then the one-nearest-neighbor classifier is built. The various experiments have shown that our proposed method has a satisfactory classification performance as compared with the current state-of-the-art approaches.

REFERENCES

[1] S. Selvan and S. Ramakrishnan, "SVD-based modeling for image texture classification using wavelet transformation," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2688–2696, Nov. 2007.
[2] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback–Leibler distance," *IEEE Trans. Image Process.*, vol. 11, no. 2, pp. 146–158, Feb. 2002.
[3] S. K. Choy and C. S. Tong, "Supervised texture classification using characteristic generalized Gaussian density," *J. Math. Imag. Vis.*, vol. 29, no. 1, pp. 35–47, Sept. 2007.
[4] L. Li, C. S. Tong, and S. K. Choy, "Texture classification using refined histogram," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1371–1378, May 2010.
[5] S. K. Choy and C. S. Tong, "Statistical properties of bit-plane probability model and its application in supervised texture classification," *IEEE Trans. Image Process.*, vol. 17, no. 8, pp. 1399–1405, Aug. 2008.
[6] M. H. Pi, C. S. Tong, S. K. Choy, and H. Zhang, "A fast and effective model for wavelet subband histograms and its application in texture image retrieval," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3078–3088, Oct. 2006.
[7] S. K. Choy and C. S. Tong, "Statistical wavelet subband characterization based on generalized Gamma density and its application in texture retrieval," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 281–289, Feb. 2010.
[8] A. Laine and J. Fan, "Texture classification by wavelet packet signatures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1186–1191, Nov. 1993.
[9] T. Randen and J. H. Husoy, "Filtering for texture classification: A comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 4, pp. 291–310, Apr. 1999.
[10] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. Image Process.*, vol. 4, no. 11, pp. 1549–1560, Nov. 1995.
[11] G. V. de Wouwer, P. Scheunders, and D. V. Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE Trans. Image Process.*, vol. 8, no. 4, pp. 592–598, Apr. 1999.
[12] S. C. Kim and T. J. Kang, "Texture classification and segmentation using wavelet packet frame and Gaussian mixture model," *Pattern Recognit.*, vol. 40, no. 4, pp. 1207–1221, Apr. 2007.
[13] X. Liu and D. L. Wang, "Texture classification using spectral histograms," *IEEE Trans. Image Process.*, vol. 12, no. 6, pp. 661–670, June 2003.
[14] [Online]. Available: [Online]. Available: http://www.ux.uis.no/tranden/brodatz.html
[15] [Online]. Available: [Online]. Available: http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html
[16] H. Le Borgne, A. G. Dugue, and N. E. O'Connor, "Learning midlevel image features for natural scene and texture classification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 3, pp. 286–297, Mar. 2007.
[17] H. Lategahn, S. Gross, T. Stehle, and T. Aach, "Texture classification by modeling joint distributions of local patterns with Gaussian mixtures," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1548–1557, Jun. 2010.
[18] Z. H. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1657–1663, Jun. 2010.
[19] Z. H. Guo, L. Zhang, and D. Zhang, "Rotation invariant texture classification using LBP variance (LBPV) with global matching," *Pattern Recognit.*, vol. 43, no. 3, pp. 706–719, Mar. 2010.