TEXTURE CLASSIFICATION USING WAVELET-DOMAIN BDIP AND BVLC FEATURES

Hyun Joo So, Mi Hye Kim, and Nam Chul Kim

School of Electrical Engnieering and Computer Science, Kyungpook National University
1370 Sankyukdong, Bukgu, 702-701, Daegu, Korea
phone: + (82) 53 950 5530, fax: + (82) 53 950 5505, email: nckim@ee.knu.ac.kr
web: vcl.knu.ac.kr

ABSTRACT

In this paper, we propose a texture classification method using local texture features BDIP (block difference of inverse probabilities) and BVLC (block variation of local correlation coefficients) in wavelet domain. BDIP and BVLC are known to be good texture features which are bounded and well normalized to reduce the effect of illumination and catch the own properties of textures effectively. In the method, a target image is first decomposed into wavelet subbands. BDIPs and BVLCs are then computed in wavelet subbands. The means and standard deviations of subband BDIPs and BVLCs and the subband standard deviations are fused into a texture feature vector. Finally, the Bayesian distance between the feature vector of a query image and that of each class is stably measured and it is classified into the class of minimum distance. Experimental results for three test databases (DBs) show the proposed method yields excellent performances.

1. INTRODUCTION

Texture classification is one of tasks essential to various image processing fields such as image retrieval, image segmentation, pattern recognition, and robot vision. In the literature, one can find numerous texture feature extraction methods which are performed in the spatial domain. A lot of them utilize gradient operators, Markov random field models, or gray-level co-occurrence probability (GLCP) to obtain texture features [1], [2]. GLCP [1] is defined as the joint probability between a quantized center pixel and one of quantized neighboring pixels. Since various neighbors are available according to the quantization scheme and their distances and directions about a center pixel, diverse GLCPs may be defined. Texture features extracted from GLCPs include energy, variance, correlation coefficient, entropy, sum entropy, difference entropy, information of correlation, and so on.

Spatial-domain texture feature extraction methods mentioned above have a weak point that an image to be classified is analyzed at a fixed scale. Many researchers have introduced Gabor filters or wavelet transform into texture classification to overcome this problem [3]-[5]. Especially, wavelet transform which allows us multiresolution analysis has been widely applied to image processing areas such as image compression and image retrieval as well as texture classification. Wavelet transform decomposes a spatial image into spatial-frequency images that are called subbands. The subbands have different resolutions according to decomposition levels, each of which has either smooth or detail information on the spatial image.

Among recent texture classification methods using wavelet transform, Selvan et al. [3] proposed a method which uses singular value decomposition (SVD) to extract the distribution parameters of singular values in the detail wavelet subbands as texture features and classify a target image by the minimum Kullback-Leibler distance classifier. Van de Wouwer et al. [4] extracted two texture feature sets, features from GLCP and coefficient distributions from the detail subbands. Wang et al. [5] adopted subband energies in wavelet packet transform domain as texture features and linear regression on pairs of subband energies for classification. Many other wavelet-based methods are also found in [6]-[8].

Besides the studies mentioned above, there have been works worthy of notice that fuse local texture features BDIP and BVLC and apply them to image retrieval [9], [10], face detection [11], ROI determination [12], and volume segmentation [13], resulting in yielding excellent performances. BDIP is a kind of nonlinear gradient operator normalized by local maximum, which is known to effectively measure local brightness variations so that edges and valleys are extracted well. BVLC is a maximal difference between local correlations according to orientations normalized by local variance, which is known to measure texture smoothness well [9]. The excellent performance of BDIP and BVLC comes from that both of them are bounded and well-normalized to reduce the effect of illumination. It seems to be well matched with the color constancy that the human visual system tries to lessen the influence of illumination to perceive a scene. Especially in [10], both of BDIP and BVLC were modified for the application on the wavelet domain, which was shown to produce a somewhat noticeable performance improvement in image retrieval.

In this paper, we propose a texture classification method using the fusion of moments of subband coefficients, subband BDIPs, and subband BVLCs. As moments, their means and variances are used which are global statistics of local features related to local statistics (local variation and correlation) in a moving window. The classical Bayesian classifier is chosen as a classifier. Section 2 explains GLCP and Haralick features, spatial-domain BDIP and BVLC, and wavelet-

domain BDIPs and BVLCs. In Section 3, the proposed classification method is described. Section 4 discusses experimental results and Section 5 shows the conclusions.

2. CONVENTIONAL TEXTURE FEATURES

In this section, we describe the conventional features which are partly used in the proposed method or in some existing methods to be compared for performance evaluation.

2.1 GLCP and Haralick Features

The estimate of joint probability required to compute Haralick features is obtained by averaging GLCPs over interested distances and orientations as follows:

$$P(i, j) = \frac{1}{K} \sum_{(d,\theta)} \Pr(i, j | d, \theta, Q) \tag{1}$$

where $P(i, j)$ denotes the estimate of joint probability and $\Pr(i, j \mid d, \theta, Q)$ a GLCP or relative frequency of a pixel pair measured in a distance d between the pair, its orientation θ, and a quantization scheme Q for the pair quantized into i and j, respectively. The parameter K is the quantity related to the probability of (d, θ, Q). In [1], Haralick et al. used (1) to compute 28 textural features such as energy, variance, correlation coefficient, entropy, sum entropy, difference entropy, information of correlation and so on.

2.2 BDIP and BVLC in the Spatial Domain

Spatial-domain BDIP for a target image I is defined as

$$BDIP(x, y) = \frac{\frac{1}{|R(x, y)|} \sum_{(p,q)\in R(x,y)} [I_{\max}(x, y) - I(p, q)]}{I_{\max}(x, y)} \tag{2}$$

where $I(x, y)$ denotes the value at a pixel (x, y) in the image I, $R(x, y)$ a local region whose center is the pixel (x, y), $|R(x, y)|$ the number of pixels in the region, and $I_{\max}(x, y)$ the maximum value in the local region. The numerator is selected as a representative gradient in the local region, which is defined by the averaged difference between the maximum pixel value and each pixel value in the local region, and the denominator as a representative value in the region, which is defined by the maximum pixel value. So, the division gives the result of gradient operator normalized by the representative, which yields a sketch-like image. It is shown in a BDIP image that object boundaries are localized well, and moreover, the intensity variation in dark regions is more emphasized [6], [7].

The computation of BVLC starts from correlation coefficients in a local region, which are defined as

$$\rho(x, y; r) = \frac{1}{\sigma(x, y)\sigma((x, y) + r)} \times \left[\frac{1}{|R(x, y)|} \times \sum_{(p,q)\in R(x,y)} I(p, q) I((p, q) + r) - \mu(x, y)\mu((x, y) + r) \right] \tag{3}$$

where r denotes a shifting orientation and $\mu(x, y)$ and $\sigma(x, y)$ are the mean and standard deviation in a local region $R(x, y)$, respectively. The terms $\mu((x, y)+r)$ and $\sigma((x, y)+r)$ are the mean and standard deviation in a local region shifted by r from (x, y), respectively.

BVLC is then defined as

$$BVLC(x, y) = \max_{r\in O_k} [\rho(x, y; r)] - \min_{r\in O_k} [\rho(x, y; r)] \tag{4}$$

where O_k denotes a set of orientations with r of distance k. For instance, O_k may be chosen as $O_k = \{(-k, 0), (0, -k), (0, k), (k, 0)\}$. The value of BVLC is determined as the difference between the maximum and minimum values of the local correlation coefficients according to orientations. The higher the degree of roughness in the local region is, the larger the value of BVLC [6], [7].

2.3 Wavelet-Domain Features

In wavelet transform or decomposition, a two-dimensional image is filtered by LPF (low pass filter) and HPF (high pass filter) along the horizontal direction and vertical direction, and so one level decomposition yields four subbands: one smooth subband (LL) and three detail subbands (LH, HL, and HH). Consecutively, the subband LL may be further decomposed in the same way. In general, each subband is downsampled by two to remove redundancies and in result the total size of four subbands becomes the same as that of the decomposition input.

Let $W_{l,b}$ be a subband image in the lth level where b denotes one of four subband, LL, LH, HL, and HH. BDIP in the wavelet domain $W_{l,b}$ is defined as [7]

$$BDIP_{l,b}(x, y) = \frac{\frac{1}{|R(x, y)|} \sum_{(p,q)\in R(x,y)} (M_{l,b}(x, y) - W_{l,b}(p, q))}{M_{l,LL}(x, y)} \tag{5}$$

where $W_{l,b}(p, q)$ denotes the wavelet coefficient at (p, q) in the subband $W_{l,b}$. $M_{l,b}(x, y)$ is the maximal value in a local region $R(x, y)$ in the subband $W_{l,b}$. Note that the denominator is given not by the corresponding subband but by the smooth subband $W_{l,LL}$ at the same level. It makes pixels on the same position in four subbands of the same level have the same representative.

The local correlation coefficient for BVLC in the wavelet subband $W_{l,b}$ is simplified from the original BDIP as [7]

$$\rho_{l,b}(x, y; r) = \frac{\frac{1}{|R(x, y)|} \sum_{(p,q)\in R(x,y)} |W_{l,b}(p, q) - W_{l,b}((p, q) + r)|}{\frac{\nu_{l,b}(x, y) + \nu_{l,b}((x, y) + r)}{2}} \tag{6}$$

where $\nu(x, y)$ denotes the mean absolute difference between pixels in a local region $R(x, y)$ and $\nu((x, y)+r)$ the corresponding value in the local region shifted by r from (x, y). The orientation set may be selected as $O_k = \{(-k, 0), (0, -k)\}$. BVLC is then computed as

$$BVLC_{l,b}(x, y) = \max_{r\in O_k} [\rho_{l,b}(x, y; r)] - \min_{r\in O_k} [\rho_{l,b}(x, y; r)] \tag{7}$$

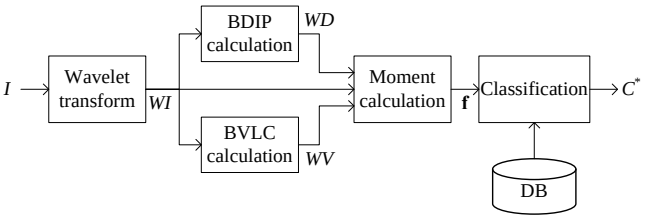

Fig. 1. Block diagram of the proposed texture classification.

3. PROPOSED TEXTURE CLASSIFICATION METHOD

In this section, we describe our classification method whose block diagram is shown in Fig. 1. An image I is first decomposed into a set of several subband images WI by wavelet transform. Next, a subband BDIP image WD and a subband BVLC image WV are obtained for each subband. Then, means and standard deviations of subbands, subband BDIPs, and subband BVLCs are calculated as texture features. Finally, the feature vector $\mathbf{f}$ is classified based on the predefined statistics of classes in a database.

In [6] and [7], BDIPs and BVLCs in the spatial and wavelet domains are extracted from 2×2 nonoverlapped regions which are partitioned from the image I or subband image WI so that the resulting size of feature images is a quarter of an original image. On the other hand, we extract them with a 3×3 moving window for every pixel so that the resulting size of feature images is equal to that of an original image. BDIP in the spatial domain is computed by (2) and BDIP in the wavelet domain by (5), respectively. However, we decide here that all BVLCs in the spatial and wavelet domains are computed according to the original definitions of (3) and (4) in the four orientations of $k = 1$.

The feature vector $\mathbf{f}$ for a target image consists of subvectors $\mathbf{f}_{l,b}$, each of which is originated from each subband $W_{l,b}$ as follows:

$$\mathbf{f}_{l,b} = [\mu_{WD_{l,b}}, \mu_{WV_{l,b}}, \sigma_{WD_{l,b}}, \sigma_{WV_{l,b}}, \sigma_{WI_{l,b}}] \tag{8}$$

where $\mu_{A_{l,b}}$ and $\sigma_{A_{l,b}}$ are mean and standard deviation of an output image $A \in \{WI, WD, WV\}$, respectively. Since the subband mean is almost zero, it is excluded in the feature vector. The feature vector is driven to the Gaussian Bayes Classifier (GBC). The distance D between a class C and a target image I of a feature vector $\mathbf{f}$ is given as

$$D(C \mid I) = (\mathbf{f} - \boldsymbol{\mu}_C)\mathbf{S}_C^{-1}(\mathbf{f} - \boldsymbol{\mu}_C)^T + \ln\left|\mathbf{S}_C\right| \tag{9}$$

where $\boldsymbol{\mu}_C$ and $\mathbf{S}_C$ denote the mean vector and the covariance matrix of trained features in the class C, respectively. The target image I is then decided as C^* which gives the minimum distance as follows:

$$C^* = \underset{C \in \{C_1, C_2, \cdots, C_N\}}{\arg\min} \; D(C \mid I) \tag{10}$$

where N denotes the number of classes.

For stabilization of distance measurement, the denominator terms of BDIP in (2), (5), and (3) are limited to $I_{\max} < 2$, $M_{l,LL} < 2$, and $\sigma^2 < 2$, respectively. Moreover, diagonal elements of the covariance matrix $\mathbf{S}_C$ in (9) are limited to 1 for wavelet feature. The values of BDIP and BVLC in (2), (5), (4), and (7) are adjusted so as to be in the range of [0, 255].

4. EXPERIMENTAL RESULTS

In this section, the performance of the proposed method is evaluated with three image DBs derived from Vistex [14] and Brodatz [15]. The first two DBs, named VS1 and BR1 here, are the same as DBs used in [4] and in [5], respectively. As mother images, VS1 has 30 Vistex images of size 512× 512 and BR1 40 Brodatz images of size 640×640, and the third DB, called BR2, shares the identical mothers of BR1 but of different size 512×512. Each mother in VS1 and BR2 consists of a class where the elements are 64 subimages of size 64×64 partitioned without overlapping from the mother image. A half of 64 subimages for each class are randomly selected for training and the other half for testing. In BR1, the mother image is partitioned into 81 subimages of size 128×128 with overlapping, and 40 subimages are for training [5].

The pair of Haar filters is chosen for wavelet transform whose decomposition level is just one without downsampling. Each of four subbands has 5 features, which results in the feature vector dimension of 20. BDIPs and BVLCs in the spatial and wavelet domains are computed at all pixels and a local region is defined as a 3×3 moving window centered at each pixel. Symmetric extension is chosen for image boundary processing. The performance of texture classification is measured as the average rate of test subimages classified correctly, which is written as

$$\frac{1}{N}\sum_{n=1}^{N}\frac{NC_n}{NT_n} \tag{11}$$

where N denotes the number of classes for each DB and NT_n and NC_n the number of test subimages and the number of correct decisions for the nth class, respectively.

For performance comparison, we also compute eight other sets of texture features and measure their performances when substituting the feature vector $\mathbf{f}$ with each of them in Fig. 1. Table 1 shows the average rates of correct decision obtained by various methods for VS1 and BR2. Haralick denotes the set of 8 Haralick features derived from GLCP: energy, variance, correlation coefficient, entropy, sum entropy, difference entropy, information of correlation, and difference mean, where the distances include $d = 1$ and 2 and the orientations $\theta = 0$, 45, 90, and 135°, and the quantization scheme Q has 256 levels. BDIP and BVLC mean the set of the global mean, standard deviation, and correlation coefficient of the spatial BDIP and the set of those of the spatial BVLC, respectively. BDIP + BVLC corresponds to the fusion of them and so its feature dimension is doubled. WI is composed of only standard deviations for four subbands. WD and WV have the global means and standard deviations of subband BDIPs and those of subband BVLCs, respectively. WD + WV stands for their fusion.

Table 1. Average rates of correct decision [%]

Feature(s) \ DB	VS1	BR2	Feature dimension
Haralick	93.1	97.3	8
BDIP	74.1	84.1	3
BVLC	60.6	75.9	3
BDIP + BVLC	92.0	98.5	6
WI	83.9	90.2	4
WD	95.7	99.1	8
WV	92.3	94.8	8
WD + WV	99.0	99.5	16
WI + WD + WV	99.0	99.9	20

In Table 1, we can see that BDIP is much better than BVLC and their fusion produces great gains of 17.9% and 14.4% for VS1 and for BR2, respectively. For BR2, WI is much worse than BDIP + BVLC. WD is also better than WV as in the spatial domain. The proposed method yields the best performance for both of VS1 and BR2. For VS1, our method gives 5.9% and 7% gains over Haralick and BDIP + BVLC, respectively. For BR2, the corresponding gains amount to 2.6%, 1.4%, and 0.4% over Haralick, BDIP + BVLC, and WD + WV, respectively. Besides, the proposed method outperforms the method in [4] of average rate 94.0% with GBC and 93.0% with KNN, where 24 histogram features and 96 GLCP Haralick features from 12 detail subbands at 4-level wavelet transform are extracted.

Since element images of BR1 are obtained with severe overlapping, they reveal high similarity to each other, we can expect that the performances of all features for BR1 are very high compared with VS1 and BR2. Actually we get correct decision of 100 % by WD + WV as well as by WI + WD + WV in cases of selecting 40 training subimages randomly. In addition, WD + WV gives 100% and WI + WD + WV 99.8% in case of selecting the upper 40 subimages for training. It can also be shown that our method excels the method in [5] of average rate 97.2%, where linear regression is performed on pairs of subband energies by 3-level wavelet packet transform. It is also shown that the stabilization of distance measurement for covariance matrices yields the performance gain of 0.1% ~ 0.4% for wavelet features.

5. CONCLUSION

A texture classification method has been proposed which uses the fusion of wavelet-domain BDIP and BVLC moments and wavelet energies. The fused features of test images were classified by the Gaussian Bayes classifier. It was shown in experiments for three test DBs that the good normalization property of BDIP and BVLC effected excellent behaviors in the wavelet domain, too. As a result, our fused feature was shown to yield the decision rates of 97.4% for some Vistex images, 99.8% for some Brodatz images, and 100% for other Brodatz images.

6. ACKNOWLEDGEMENT

This work was supported by the Brain Korea 21 Project (BK21).

REFERENCES

[1] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Sys., Man, Cybern.*, vol. SMC-3, no. 6, Nov. 1973.

[2] M. Tuceryan and A. K. Jain, *Texture Analysis, In*: *The Handbook of Pattern Recognition and Computer Vision*, World Scientific, 2nd ed., New Jersey, 1998.

[3] S. Selvan and S. Ramarkrishnan, "SVD-based modeling for image texture classification using wavelet transformation," *IEEE Trans. IP*, vol. 16, no. 11, Nov. 2007.

[4] G. Van de Wouwer, P. Scheunders, and D. Van Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE Trans. Image Process.*, vol. 8, no. 4, Apr. 1999.

[5] Z. Z. Wang and J. H. Yong, "Texture analysis and classification with linear regression model based on wavelet transform," *IEEE Trans. Image Process.*, vol. 17, no. 8, pp. 1421-1430, Aug. 2008.

[6] D. A. Clausi and H. Deng, "Design-based texture feature fusion using Gabor filters and co-occurrence probabilities," *IEEE Trans. Image Process.*, vol. 14, no. 7, Jul. 2005.

[7] P. S. Hiremath and S. Shivashankar, "Wavelet based co-occurrence histogram features for texture classification with an application to script identification in a document image," *Pattern Recognition Letters*, vol. 29, no. 9, pp. 1182-1189, Jul. 2008.

[8] E. Avci, A. Sengur, and D. Hanbay, "An optimum feature extraction method for texture classification,", *Expert Systems with Applications*, vol. 36, no. 3, part 2, pp. 6036-6043, Apr. 2009.

[9] Y. D. Chun, N. C. Kim, and I. H. Jang, "Content-based image retrieval using multiresolution color and texture features," *IEEE Trans. Multimedia*, vol. 10, no. 6, pp. 1073-1084, Oct. 2008.

[10] Y. D. Chun, S. Y. Seo, and N. C. Kim, "Image retrieval using BDIP and BVLC moments," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 9, pp. 951-957, Sep. 2003.

[11] H. J. So, M. H. Kim, Y. S. Chung, and N. C. Kim, "Face detection using sketch operators and vertical symmetry," *FQAS 2006, Lecture Notes in Artificial Intelligence*, vol. 4027, pp. 541-551, Jun. 2006.

[12] T. D. Nguyen, S. H. Kim, and N. C. Kim, "An automatic body ROI determination for 3D visualization of a fetal ultrasound volume," *KES 2005, Lecture Notes in Artificial Intelligence*, vol. 3682, no. 2, pp. 145-153, Sep. 2005.

[13] J. I. Kwak, M. N. Kim, S. H. Kim, and N. C. Kim, "3D segmentation of breast tumor in ultrasound images," in *Proc. SPIE's Medical Imaging*, vol. 5203, San Diego, USA, pp. 193-200, Feb. 15-20. 2003.

[14] [Online]. Available: http://www.media.mit.edu/vismod

[15] [Online]. Available: http://www.ux.uis.no/~tranden/brodatz.html