JPEG2000-Based Image Features with Its Application to Texture Segmentation

Hsi-Chin Hsin* and Tze-Yun Sung†
*National United University, Miao-Li 36003, Taiwan
E-mail: hsin@nuu.edu.tw Tel: +886-37-381619
†Chung Hua University, Hsinchu City 300-12, Taiwan

***Abstract*— In this paper, a context-based wavelet histogram (CBWH) is proposed to characterize image textures. A simple method is also developed to estimate bit-plane probabilities of the CBWH, which can be used as features to segment textured images. Based on the arithmetic coder of the JPEG2000 standard, the CBWH-derived image features can be obtained directly from the MQ table. As a result, images can be segmented in the JPEG2000 domain. The potential of this new scheme is shown by experimental results.**

I. INTRODUCTION

Texture segmentation is important in many applications, raging from industrial monitoring to medical diagnosis. Early work on feature extraction is mainly focused on the analysis of gray levels at one single scale [1]-[3]. As observed in the human visual system (HVS), the received image is decomposed into a collection of band-pass sub-images and then analyzed by the simple visual cortical cells, which can be well modeled by a set of Gabor filters with orientations and spatial frequencies properly tuned [4]; various tuning algorithms by global Fourier analysis [5] or local Fourier analysis [6] have been proposed. Wavelet theory provides a solid framework for representing images at multiple scales [7]. Mallat applied wavelet transform with an efficient pyramid structured algorithm to image analysis [8]. In order to extend the bases functions and overcome the shift variant problem of wavelet transform, Pun proposed an adaptive wavelet packet transform for extracting shift invariant texture features [9].

In many cases, image compression is necessary. The widely used discrete cosine transform (DCT) based image coder known as Joint Photographic Expert Group (JPEG) shows satisfactory results [10]. To improve the overall compression performance with additional advantages, e.g., progressive transmission and embedded coding, various discrete wavelet transform (DWT) based image coders have been proposed [11]-[13]. Image coding with progressive transmission is particularly desirable for the Internet streaming and database browsing. By embedded coding, the original image can be encoded into a single code stream, and based on which the decoded images at various bit rates can be obtained; the embedded code stream of an image is usually organized in decreasing order of information importance. The JPEG2000 standard adopts DWT as the underlying transform algorithm [14]. After DWT, the higher detail information of an image is projected onto the shorter basis function with higher spatial resolution; the lower detail information is projected onto the larger basis function with higher spectral resolution; this property matches the characteristics of HVS [15]. For image retrieval in a large JPEG2000 database, Pi proposed a scheme to extract features from the compressed image rather than from the wavelet coefficients or the pixel gray levels [16]; in which, the bit-plane probabilities obtained by counting the number of 1-bits were used to model the wavelet histograms.

For image segmentation, we propose a new scheme to extract the local signatures of the wavelet histogram from a JPEG2000 code stream. The remainder of this paper proceeds as follows. In Section II, wavelet transform and the JPEG2000 standard are reviewed briefly. In Section III, a context-based wavelet histogram is proposed to characterize images' signatures; the corresponding bit-plane probabilities that can be obtained directly from the MQ table defined in JPEG2000 are used as texture features for segmenting images in the JPEG2000 domain. Experimental results are presented in Section IV. Conclusion is given in Section V.

II. REVIEW OF WAVELET TRANSFORM AND JPEG2000

A. Wavelet Transform

Wavelet transform (WT) provides a multi-resolution representation with various desirable properties, e.g. subband decomposition with orientation selectivity, joint space-spatial frequency localization, self similarity of wavelet coefficients across subbands of the same orientation, and energy clustering within each subband. For image applications, 2-D WT can be obtained by using the tensor product of 1-D WT. Fig. 1 shows a 3-level 2-D WT, where HL_ℓ, LH_ℓ, and HH_ℓ denote wavelet subbands of coefficients: $D_\ell^1(m,n)$, $D_\ell^2(m,n)$, and $D_\ell^3(m,n)$ representing the detail information at resolution ℓ in the horizontal, vertical, and diagonal orientations, respectively; LL_3 denotes the approximation at the coarsest resolution 3. The original image is usually taken as the scaling coefficients $S_0(m,n)$ at the finest resolution 0. After 1-level WT, $S_0(m,n)$ is decomposed into $S_1(m,n)$, $D_1^1(m,n)$, $D_1^2(m,n)$, and $D_1^3(m,n)$; $S_0(m,n)$ can be reconstructed from $S_1(m,n)$, $D_1^1(m,n)$, $D_1^2(m,n)$, and $D_1^3(m,n)$ by the inverse wavelet transform.

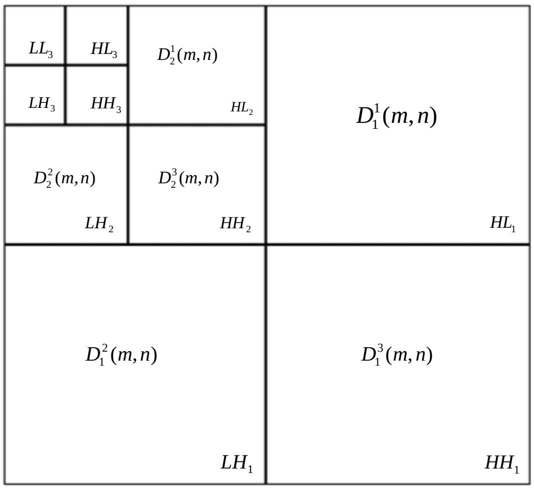


Fig. 1 3-level 2-D DWT Example

B. The JPEG2000 Standard

The core of JPEG2000, known as the embedded block coding with optimized truncation (EBCOT) algorithm, is to exploit the energy clustering property of wavelet coefficients. EBCOT is a two-tier algorithm: tier-1 performs bit-plane coding (BPC) followed by arithmetic coding (AC); tier-2 aims for rate distortion optimization [13]-[14]. The quantized wavelet coefficients of an image are coded by a context-based arithmetic coder known as the MQ coder, in which probability models are stored in the MQ table. Fig. 2(a) and Fig. 2(b) depict block diagrams of the JPEG2000 encoder and decoder, respectively.

In EBCOT, BPC is divided into three coding passes, namely the significance propagation pass, the magnitude refinement pass, and the clean-up pass. Four primitive coding operations: the significance coding operation, the sign coding operation, the magnitude refinement coding operation, and the cleanup coding operation are performed in the corresponding coding passes. For a wavelet coefficient that is currently insignificant, if any of the 8 neighboring wavelet coefficients are already significant, it is coded in the significance propagation pass using the significance coding operation; otherwise, it is coded in the cleanup pass using the cleanup coding operation; if this coefficient becomes significant, its sign is coded immediately using the sign coding operation. In the magnitude refinement pass, magnitudes of the significant wavelet coefficients found in the previous coding passes are updated using the magnitude refinement coding operation.

The output bit streams of the coding passes can be further coded by an arithmetic coding engine to improve the compression performance. Based on the current state of the 8 neighboring coefficients, the context-based MQ coder defines 18 context labels, and stores their respective probability modes in the MQ table. More precisely, 10 context labels are defined for the significance coding operation and the cleanup coding operation, 5 context labels for the sign coding operation, and 3 context labels for the magnitude refinement coding operation.

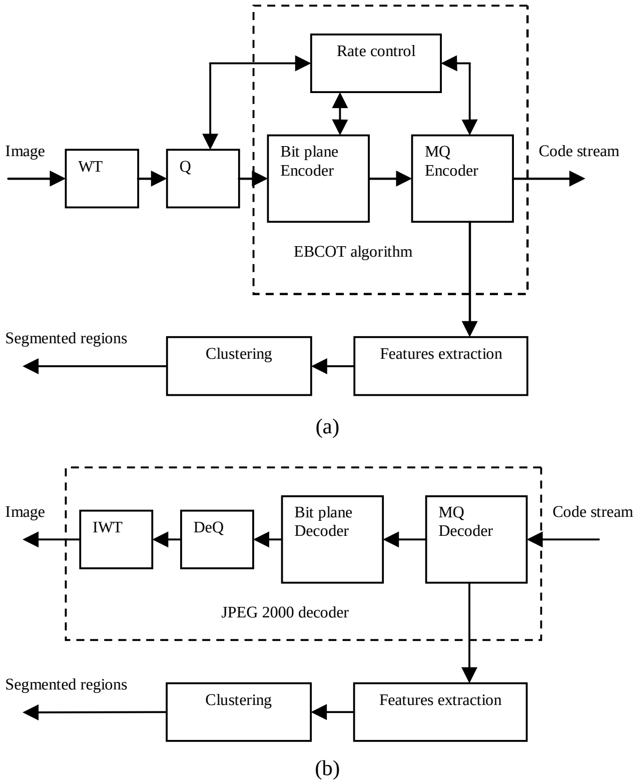


Fig. 2 Block diagram of JPEG2000 (with features extracted for image segmentation) (a) encoder (b) decoder

III. Image Features Extracted directly From JPEG2000 Code Streams

In this section, the JPEG2000-based features are proposed to characterize the underlying wavelet histograms of an image; a simple, effective method is also developed to estimate these features. As a result, images can be segmented by clustering the proposed features in the JPEG2000 domain.

A. Context-Based Wavelet Histograms

As noted in [8], distribution of wavelet coefficients is typically symmetrical about zero, and therefore the histogram of the absolute wavelet coefficients can be effectively used to characterize image textures. In order to reduce the dimension, Do and Vetterli utilized the generalized Gaussian distribution (GGD) to model wavelet histograms [17]; they used the GGD parameters as features for texture retrieval.

In JPEG2000, wavelet coefficients are quantized into n bit-planes. It is likely that variables on each bit-plane, which are either 1 or 0, are independent across bit-planes, and the joint probability mass function (PMF) representing the wavelet histogram can be written by

$$P(|c| = x) = \prod_{i=0}^{n-1} P_i(x_i) \tag{1}$$

where c is a wavelet coefficient, $x = \sum_{i=0}^{n-1} x_i \cdot 2^i;\ x_i \in \{0,1\}$, and $P_i(\circ)$ is the i^{th} bit-plane PMF. Thus, Pi et al. utilized the set of PMF: $\{P_i; i = 0, \cdots, n-1\}$ to represent wavelet histograms [16]; in which, $P_i(x_i = 1)$ was estimated by counting the number of 1-bits on the i^{th} bit-plane for each wavelet subband. Though counting operation is efficient in comparison with the estimation of GGD parameters, the local PMF is still needed for image segmentation. In the next subsection, we propose a simple method to estimate the local PMF based on the MQ table of JPEG2000, which requires no extra computation at all.

B. References

In JPEG2000, a big image can be divided into sub-images called tiles; each tile is decomposed into subbands by DWT; each subband is partitioned into small blocks called code blocks; each quantized code block is coded independently using the EBCOT algorithm, from the most significant bit-plane to the least significant bit-plane. One of the most important features of wavelet transform is the energy clustering of wavelet coefficients within each subband. The following interesting questions are thus raised. 1) Is it possible that image segmentation can be performed directly in the JPEG2000 domain such that the burden of decoding computation can be avoided? 2) Is there a common piece of information, based on which features can be constructed at both encoder and decoder? If so, there is no need to transmit these features from encoder to decoder. 3) For computational simplicity, are there any easy ways to solve these questions?

Recall that EBCOT is a two-tier algorithm. In tier-1, probabilities of the next more probable symbol (NMPS) and the next less probable symbol (NLPS) are stored in the MQ table, which is available at both encoder and decoder. If the set of PMF representing wavelet histograms can be obtained from the MQ table, the above-mentioned questions 1 and 2 can be solved simultaneously. Fig. 2(a) and Fig. 2(b) show the JPEG2000 encoder and decoder with features extracted from the MQ table, respectively. Thereafter, image segmentation can be obtained by clustering these features.

Motivated by the MQ coder adopted by JPEG2000, where the current LPS probability is stored in the MQ table, a simple method is thus proposed to estimate the local PMF using the following rule:

$$P_i(x_i = 1) = \begin{cases} Qe\_Value & \text{if MPS} = 0 \\ 1 - Qe\_Value & \text{if MPS} = 1 \end{cases} \tag{2}$$

where $P_i(x_i = 1)$ is the probability of 1-bit for variable x_i on the i^{th} bit-plane, Qe_Value is the LPS probability stored in the MQ table. Fig. 3 depicts flow chart of the proposed method in the framework of JPEG2000. It is favorably noted that the proposed method is simple.

Take the test image shown in Fig. 5(a) as an example. For the HH_2 subband, Fig. 4 shows the estimated probabilities of 1-bit from the most significant bit-plane to the least significant bit-plane. Where, the horizontal and vertical axes are the bit-plane index and the probability value, respectively. Note that the average of the local PMF obtained by the proposed method (plus symbols) is consistent with the one obtained by counting the number of 1-bits (circle symbols) [16]. It demonstrates that the JPEG2000-based local PMF are suitable for image segmentation.

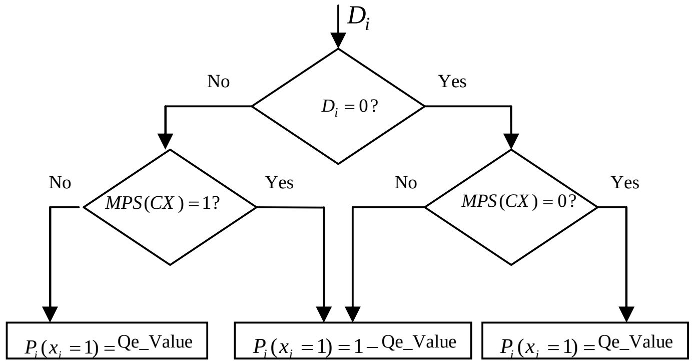


Fig. 3 Flow chart of the proposed method to estimate the local PMF: $P_i(x_i = 1)$ on the i^{th} bit-plane; *CX* denotes the context of the input: D_i

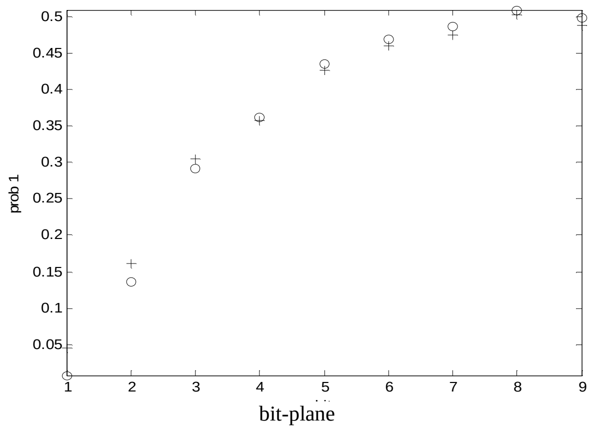
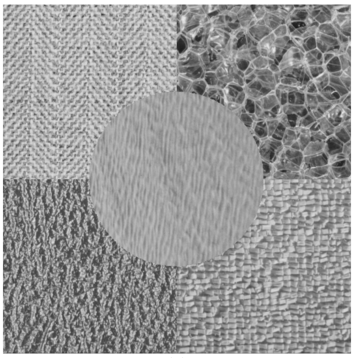


Fig. 4 Estimated probabilities of 1-bit for the HH_2 subband of Fig. 5(a), from the most significant bit-plane to the least significant bit-plane, by counting the number of 1-bits (circle symbols) and by using the proposed method (plus symbols)

IV. EXPERIMENTAL RESULTS

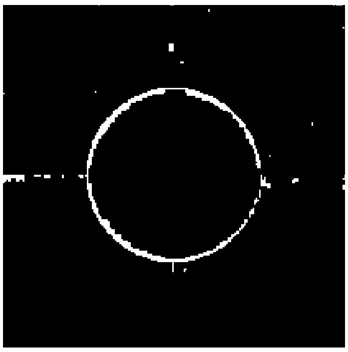
The JPEG2000-based texture features that characterize the context-based wavelet histogram are evaluated on grayscale textured images. All the texture mosaic images are produced from the Brodatz textures [18]. Features obtained from the MQ table are filtered by using a 3 x 3 low-pass Gaussian window to construct feature vectors. The K-means algorithm is used to cluster these feature vectors for the segmentation task. The bi-orthogonal wavelet with 9/7-filters is used.

The first two texture mosaic images are of size 512 x 512, Figure 5(a) shows the first test image, which is composed of five Brodatz textures, namely herringbone weave, pressed calf leather, raffia, plastic bubbles, and water. The segmentation result and the error image with white pixels representing misclassified pixels are shown in Figures 5(b) and 5(c), respectively. Figure 6(a) shows the second test image, which is composed of four Brodatz textures, namely woolen cloth, water, grass, and sand. Figures 6(b) and 6(c) show the segmentation result and the error image, respectively, by using the proposed method. As expected, the MQ table provided by the JPEG2000 standard is very effective to estimate the local PMF of wavelet coefficients.



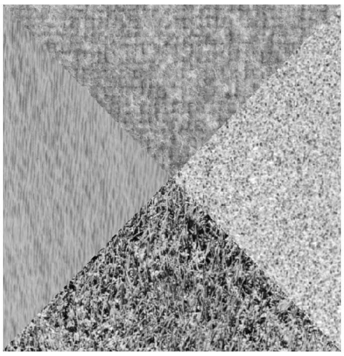
(a)

(b)

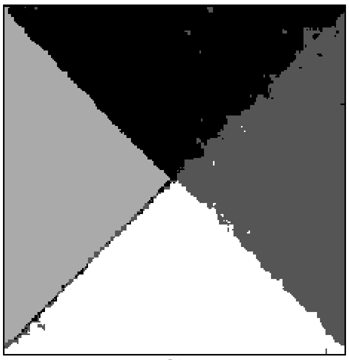


(c)

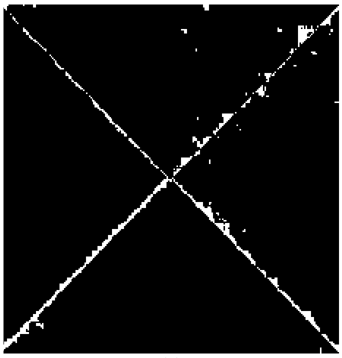
Fig. 5 (a) Mosaic 1 (b) segmentation result (c) error image



(a)



(b)



(c)

Fig. 6 (a) Mosaic 2 (b) segmentation result (c) error image

For comparison with the complete algorithm [19], two more texture mosaic images of 256 x 256 pixels are presented. Test image 3 shown in Fig. 7(a) is composed of five Brodatz textures, namely woolen cloth, sand, straw, water, and bark; Test image 4 shown in Fig. 7(b) is composed of two Brodatz textures: wood and grass. Their respective segmentation results and error images by the proposed method are shown in Fig. 7(c) and Fig. 7(d), Fig. 7(e) and Fig. 7(f). It is noted that the misclassified pixels are likely to occur near the boundaries between different textures regions, where the content of the MQ table is changed rapidly. Nevertheless, the pixel classification errors given in Table 1 show that the proposed method is almost comparable with the complete algorithm.

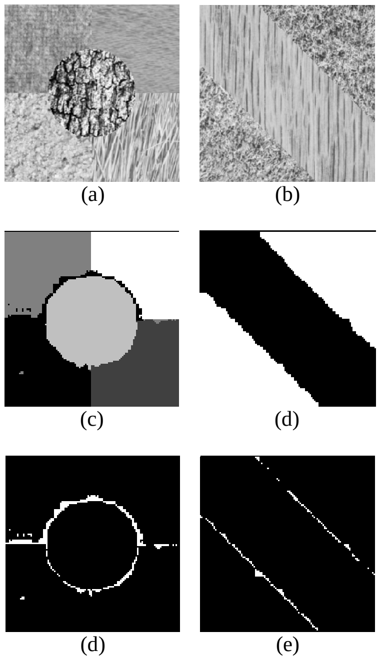


Fig. 7 (a) ~ (b) Test images 3 and 4 (c) ~ (d) segmentation results by the proposed method (e) ~ (f) error image

TABLE 1: COMPARISON OF PIXEL CLASSIFICATION ERRORS

	The complete algorithm [19]	The proposed method
Test image 3	3.3%	3.5%
Test image 4	1.0%	1.4%

The probability Rand index (PRI) used as a measure of similarity in computer vision applications [20] is also considered here. It takes values in [0, 1] with larger values meaning higher degrees of consistency in segmentation results. By comparison with the ground truth segmentation, the PRI performance of the proposed method is given in Table 2. As noted, a high degree of PRI similarity is indicative of the potential of the proposed segmentation scheme in the JPEG2000 domain.

TABLE 2: PRI PERFORMANCE OF THE PROPOSED METHOD

	PRI
Test image 1	0.9828
Test image 2	0.9732
Test image 3	0.9721
Test image 4	0.9712

Figure 8(a) shows the last test image of 320 x 320 pixels taken from the Berkeley Segmentation Dataset and Benchmark [21]. Segmentation was done by clustering the JPEG2000-based features into three clusters. Figure 8(b) shows the segmentation result. The texture boundary of zebras was quite well localized as shown in Fig. 8(c)

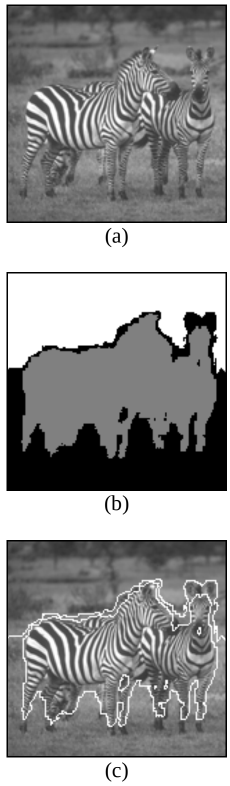


Fig. 8 (a) Test image 5 (b) segmentation result (c) original image with overlay of the segmented texture boundary

V. CONCLUSIONS

Wavelet transform has drawn a lot of attention to image processing applications. The wavelet-based multiresolution representation matches the characteristics of HVS; this property is beneficial especially for the image segmentation task. This paper proposes a new, context-based wavelet histogram in the framework of JPEG2000. Based on the MQ table defined in JPEG2000, a simple scheme has been developed to estimate bit-plane probabilities of the context-based wavelet histogram, which can be used as texture features. Since the MQ table is available at both encoder and decoder, no overhead transmission is needed for estimating the above-mentioned features from a JPEG2000 code stream. This new approach to image segmentation in the JPGE2000 domain has been evaluated on mosaic images with Bodatz textures. As expected, experimental results show that the proposed JPEG2000-based features, which can be extracted directly form the MQ table without computation, are impressively effective.

ACKNOWLEDGMENT

The authors are grateful to the maintainers of the Berkeley Segmentation Dataset and Benchmark for the PRI code. The National Science Council of Taiwan, under Grant NSC-97-2221-E-239-031-MY2 supported this work.

REFERENCES

[1] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Texture features for image classification," IEEE Trans. SMC, vol. 3, pp. 610-621, 1973.
[2] R. W. Conners and C. A. Harlow, "A theoretical comparison of texture algorithms," IEEE Trans. PAMI, vol. 2, pp. 204-222, 1980.
[3] G. C. Cross and A. K. Jain, "Markov random field texture models," IEEE Trans. PAMI, vol. 5, pp. 25-39, 1983.
[4] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two dimensional visual cortical filters," J. Opt. Soc. Amer., vol. 2, pp. 1160-1169, 1985.
[5] T. N. Tan and A. G. Constantinides, "Texture analysis based on a human visual model," Proc. IEEE ICASSP, pp. 2091-2110, 1990.
[6] O. Pichler et al., "An unsupervised texture segmentation algorithm with feature space reduction and knowledge feedback," IEEE Trans. IP, vol. 7, pp. 53-61, 1998.
[7] I. Daubechies, "Orthonormal bases of compactly supported wavelets," Commun. Pure Appl. Math., vol. 41, pp. 909-996, 1988.
[8] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans. PAMI, vol. 11, pp. 674-693, July 1989.
[9] C. M. Pun and M. C. Lee, "Extraction of shift invariant wavelet features for classification of images with different sizes," IEEE Trans. PAMI, vol. 26, pp. 1228-1233, 2004.
[10] K. R. Rao and J. J. Hwang, Techniques and Standards for Image, Video and Audio Coding, Prentice Hall, 1996.
[11] N. Sprljan, S. Grgic, and M. Grgic, "Modified SPIHT Algorithm for Wavelet Packet Image Coding," Real-Time Imaging, vol. 11, pp. 378-388, 2005.
[12] T.-Y. Sung and H.-C. Hsin, "An Efficient Rearrangement of Wavelet Packet Coefficients for Embedded Image Coding Based on SPIHT Algorithm," IEICE Trans. Fundamentals, vol. E90-A, pp. 2014-2020, 2007.
[13] D. Taubman, "High performance scalable image compression with EBCOT," IEEE Trans. IP, vol. 9, pp. 1158-1170, July, 2000.
[14] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG2000 still image compression standard," IEEE Signal Process. Mag., vol. 18, pp. 36-58, 2001.
[15] G. Strang and T. Nguyen, Wavelets and Filter Banks. Wellesley, MA: Wellesley-Cambridge, 1996.
[16] 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. IP, vol. 15, pp. 3078-3088, 2006.
[17] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," IEEE Trans. IP, vol. 11, no. 2, pp. 146-158, February 2002.
[18] P. Brodatz, Textures: A Photographic Album for Artists & Designers. New York: Dover, 1966.
[19] J. Sun, D. Gu, S. Zhang and Y. Chen, "Hidden Markov Bayesian texture segmentation using complex wavelet transform," IEE Proc.-Vis. Image Signal Process, vol. 151, no. 3, pp. 215-223, June 2004.
[20] R. Unnikrishnan, C. Pantofaru, and M. Hebert, "Toward Objective Evaluation of Image Segmentation Algorithms," IEEE Trans. PAMI, vol. 29, no. 6, pp. 929-944, June 2007.
[21] http://www.eecs.berkeley.edu/~yang/software/lossy_segmentation/