

Research Article

Adaptive Binary Arithmetic Coder-Based Image Feature and Segmentation in the Compressed Domain

**Hsi-Chin Hsin,¹ Tze-Yun Sung,² Yaw-Shih Shieh,²
and Carlo Cattani³**

¹ *Department of Computer Science and Information Engineering, National United University,
Miaoli 36003, Taiwan*

² *Department of Electronics Engineering, Chung Hua University, Hsinchu City 30012, Taiwan*

³ *Department of Mathematics, University of Salerno, Via Ponte Don Melillo, 84084 Fisciano, Italy*

Correspondence should be addressed to Tze-Yun Sung, bobsung@chu.edu.tw

Received 31 August 2011; Accepted 23 September 2011

Academic Editor: Ming Li

Copyright © 2012 Hsi-Chin Hsin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Image compression is necessary in various applications, especially for efficient transmission over a band-limited channel. It is thus desirable to be able to segment an image in the compressed domain directly such that the burden of decompressing computation can be avoided. Motivated by the adaptive binary arithmetic coder (MQ coder) of JPEG2000, we propose an efficient scheme to segment the feature vectors that are extracted from the code stream of an image. We modify the Compression-based Texture Merging (CTM) algorithm to alleviate the influence of overmerging problem by making use of the rate distortion information. Experimental results show that the MQ coder-based image segmentation is preferable in terms of the boundary displacement error (BDE) measure. It has the advantage of saving computational cost as the segmentation results even at low rates of bits per pixel (bpp) are satisfactory.

1. Introduction

Image segmentation is important in many applications, ranging from industrial monitoring to medical diagnosis. Among numerous techniques, the feature-based approach has received a lot of attention due largely to its computational efficiency [1]. However, the segmentation result is dependent on the selection of feature vectors [2–7]. Early research work on feature extraction is mainly at a single scale. It is noted that an image is decomposed into band-pass subimages by simple visual cortical cells in the human visual system (HVS) [8], which can be modeled by Gabor filters with spatial frequencies and orientations properly tuned

[9]. Wavelet transform (WT) provides an efficient multiresolution representation, in which the higher detail information of an image is projected onto the shorter basis function with higher spatial resolution, and the lower detail information is projected onto the larger basis function with higher spectral resolution. This property matches the characteristics of HVS [10]. Various WT-based schemes were proposed to extract image features at multiple scales [11–14]. In addition, the advantage of WT is to take account of the phenomena of multiscales [15–17], which is fundamental in nonlinear time series [18–20] and fractal time series [21].

With the rapid growth of multimedia technologies [22–25] and the Internet applications, image compression is still in great demand [26]. As one can see, it is desirable to extract image features in the compressed domain directly, such that the burden of decompressing an image can be avoided [14, 27, 28]. The Joint Photographic Expert Group (JPEG) standard shows satisfactory results at moderate compression rates. The JPEG2000 standard, which adopts WT as the underlying transform, is preferable for additional advantages, for example, embedded coding and progressive transmission [29, 30]. In embedded coding, the original image is coded into a single code stream, from which the decoded image at any bit rate can be obtained. For progressive transmission, which is especially beneficial to the image browsing and Internet streaming applications, JPEG2000 uses the postcompression rate distortion (PCRD) algorithm to arrange the code stream of an image in decreasing order of information importance [31]. It is based on the rate distortion theory; more specifically, the rate distortion slope (RDS) should be nonincreasing as the number of coding bits increases. For the image segmentation applications, two interesting questions are thus raised. (1) Is it possible that image segmentation can be carried out in the compressed domain such that the burden of decoding computations can be avoided? (2) Is there a common piece of information, based on which image features can be constructed at both encoder and decoder? If so, there is no need to transmit these features from encoder to decoder.

This paper presents an efficient scheme to segment an image in the compressed domain. It is a two-step algorithm. In the first step, the MQ coder-based image features are coarsely clustered into small regions known as superpixels by using the simple K -means algorithm. The inherently oversegmented superpixels are merged recursively by using the Compression-based Texture Merging (CTM) algorithm [32]. In order to avoid over merging, we propose a simple RDS-based method to terminate CTM accordingly. The remainder of this paper proceeds as follows. In Section 2, the JPEG2000 standard is briefly reviewed. In Section 3, the modified CTM algorithm with the MQ coder-based image features is proposed to segment JPEG2000 images. Experimental results are presented in Section 4. Conclusion is given in Section 5.

2. Introduction to JPEG2000

The core of JPEG2000 is the embedded block coding with optimized truncation (EBCOT) algorithm [29], which adopts wavelet transform (WT) as the underlying method for subband decompositions. WT provides many desirable properties, for example, joint space-spatial frequency localization with orientation selectivity, self-similarity of wavelet coefficients across subbands of the same orientation, and energy clustering within each subband [10]. Among various WT-based image features, the commonly used are magnitude, energy, the generalized Gaussian distribution signature, and the cooccurrence measures [11–14].

EBCOT is a two-tier algorithm. Tier-1 consists of bit-plane coding (BPC) followed by arithmetic coding (AC). Tier-2 aims for optimal rate control. In BPC, three coding passes,

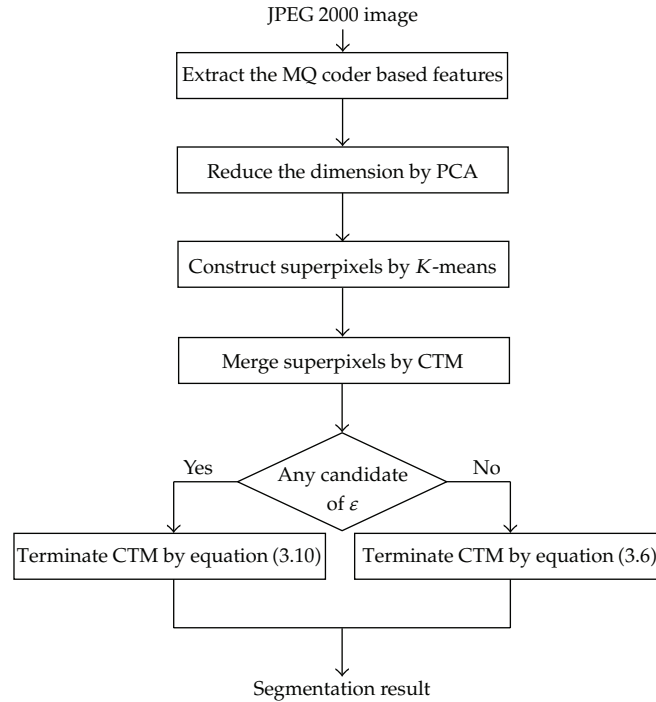


Figure 1: Flowchart of the proposed algorithm.

namely, the significance propagation (SP) pass, the magnitude refinement (MR) pass, and the clean up (CU) pass, are involved with four primitive coding operations, namely, the significance coding operation, the sign coding operation, the magnitude refinement coding operation, and the cleanup coding operation. For a wavelet coefficient that is currently insignificant, if any of the 8 neighboring coefficients are already significant, it is coded in the SP pass using the significance coding operation; otherwise, it is coded in the CU pass using the cleanup coding operation. If this coefficient becomes significant, the sign is coded immediately using the sign coding operation. In the MR pass, magnitudes of the significant coefficients are updated using the magnitude refinement coding operation. The output bit streams of coding passes can be further coded by using a context-based arithmetic coder known as the MQ coder to improve the compression performance. Based on the 8 neighboring coefficients, the MQ coder defines 18 context labels with their respective probability modes stored in the MQ table [29].

In JPEG2000, a large image can be partitioned into nonoverlapped subimages called tiles, each tile is decomposed into subbands by WT, each subband is divided into small blocks called code blocks, and each code block is independently coded from the most significant bit-plane to the least significant bit-plane. For optimal rate control, JPEG2000 adopts the postcompression rate distortion (PCRD) algorithm. Specifically, let $\{B_i\}$ be the code blocks of an image. The embedded code stream of B_i can be terminated at some point, say n_i , with a bit rate denoted by $R_i^{n_i}$; all the end points of coding passes are possible truncation points. PCRD selects the optimal truncation points to minimize the overall distortion: $D = \sum_i D_i^{n_i}$ subject to the rate constraint: $R = \sum_i R_i^{n_i} \leq R_c$, where $D_i^{n_i}$ denotes the distortion incurred by discarding the coding passes after n_i , and R_c is the target bit rate. It is noted that the

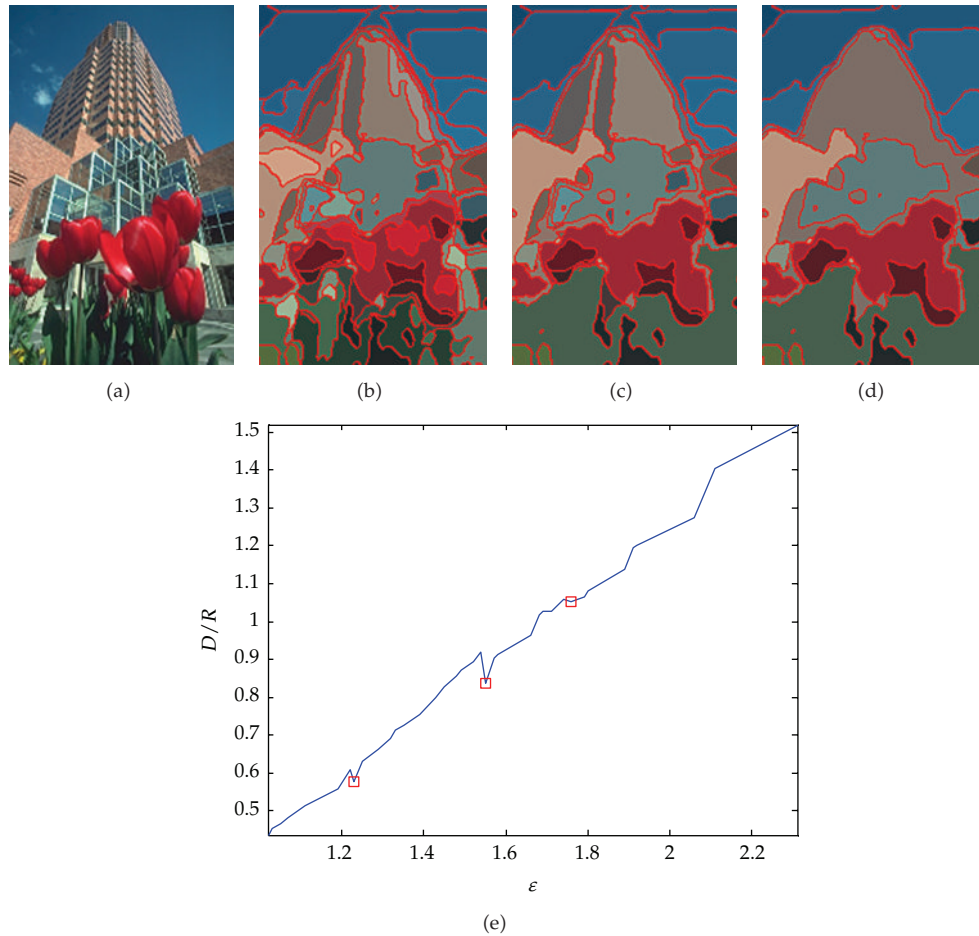


Figure 2: (a) Original image. (b) ~ (d) Segmentation results with the first, second, and third candidates of ϵ . (e) Plot of the candidates of ϵ .

coding passes with nonincreasing rate distortion slopes (RDS) are candidates for the optimal truncation points. Based on the above, we propose an efficient scheme to segment JPEG2000 images in the following section.

3. Image Segmentation in the JPEG2000 Domain

In this section, we modify the Compression-based Texture Merging (CTM) algorithm [32] to segment the MQ coder-based image features [28] in an adaptive manner. As a result, the image segmentation task can be conducted in the JPEG2000 domain directly, and the burden of decompressing computation can be avoided.

3.1. The MQ Coder-Based Image Feature

The distribution of wavelet coefficients known as the wavelet histogram has been widely used for image segmentation. As the binary variables of wavelet coefficients are almost

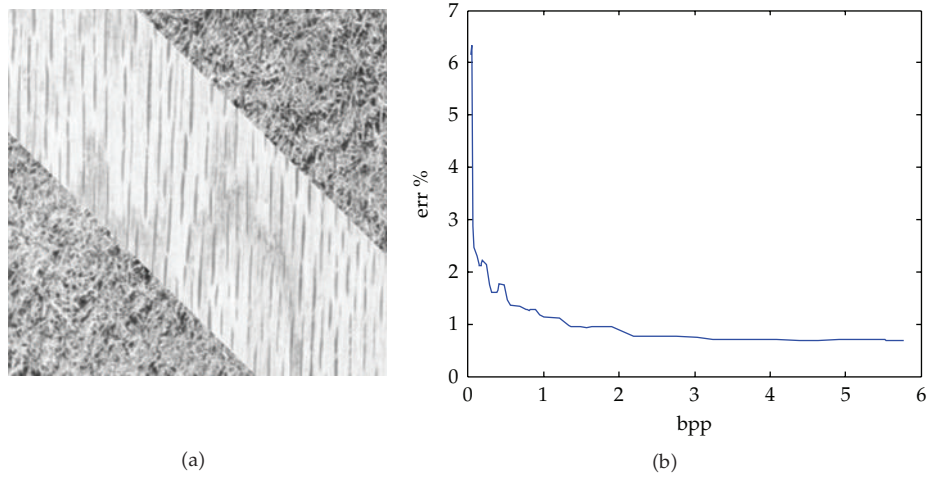


Figure 3: Performance of the proposed algorithm at various bpp rates. (a) Test image. (b) Error rates in percentage at various bpp rates.

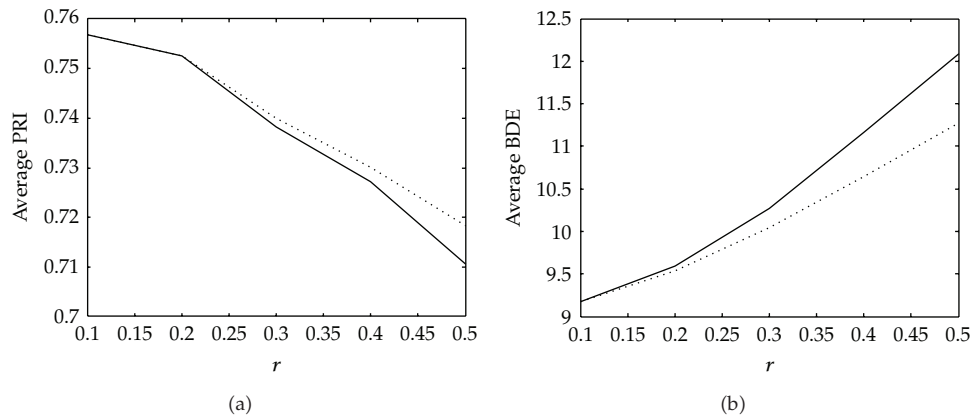


Figure 4: Average (a) PRI and (b) BDE using CTM (solid line) and the modified CTM with the MQ coder-based image features (dotted line).

independent across bit-planes [14], the joint probability mass function (PMF) representing the wavelet histogram can be approximated as

$$P(|c| = x) = \prod_{i=0}^{n-1} P_i(x_i), \quad (3.1)$$

where x is the absolute value of a wavelet coefficient, c , which can be written by

$$x = \sum_{i=0}^{n-1} x_i \cdot 2^i, \quad x_i \in \{0, 1\}, \quad (3.2)$$

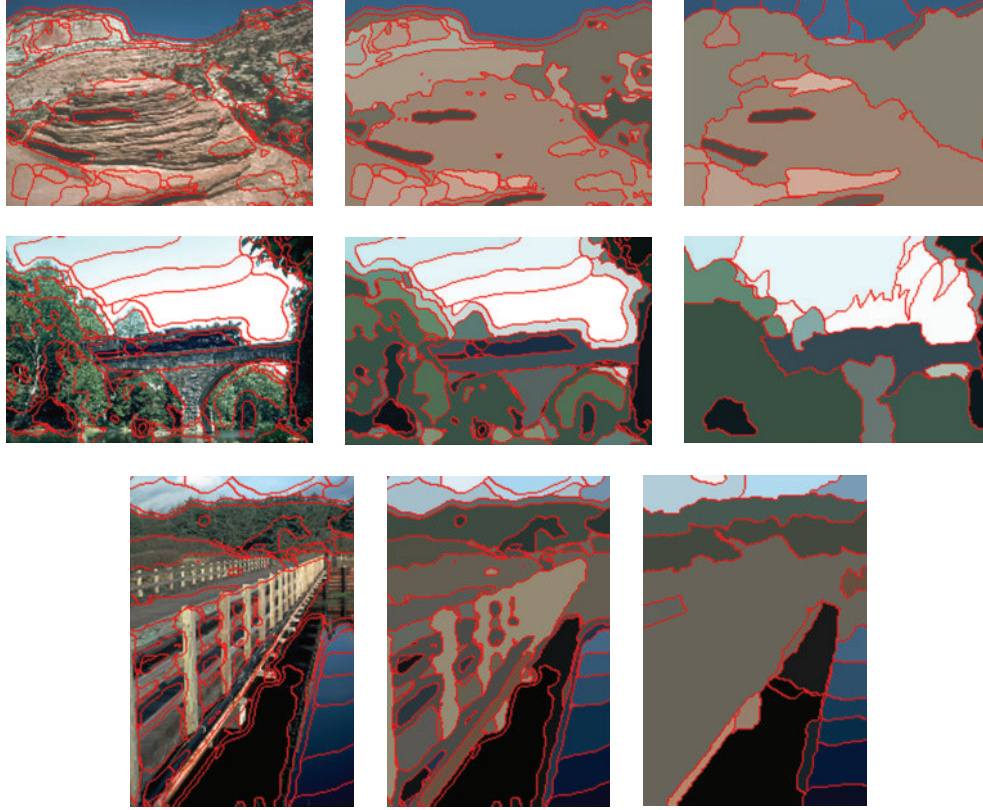


Figure 5: Landscape images: left column: original images with the boundaries obtained by using the proposed algorithm; middle and right columns: the segmentation results obtained by using the proposed algorithm and the CTM algorithm, respectively.

n is the number of bit-planes, and $P_i(\circ)$ is the i th bit-plane's PMF. Based on the MQ table defined in JPEG2000, we proposed a simple scheme to estimate the local PMF [28]. Specifically, let $P_i(x_i = 1)$ be the probability of 1 bit for variable x_i on the i th bit-plane, which can be obtained from the MQ table as follows:

$$P_i(x_i = 1) = \begin{cases} \text{Qe_Value} & \text{if MPS} = 0, \\ 1 - \text{Qe_Value} & \text{if MPS} = 1, \end{cases} \quad (3.3)$$

where Qe_Value is the probability of less probable symbol (LPS) stored in the MQ table, and MPS stands for more probable symbol. Note that the set $\{P_i(x_i = 1); i = 0, \dots, n-1\}$ obtained from the MQ table can be used to estimate the local PMF. As the MQ table is available at both encoder and decoder, there is no need to transmit the overhead information to construct the MQ features.

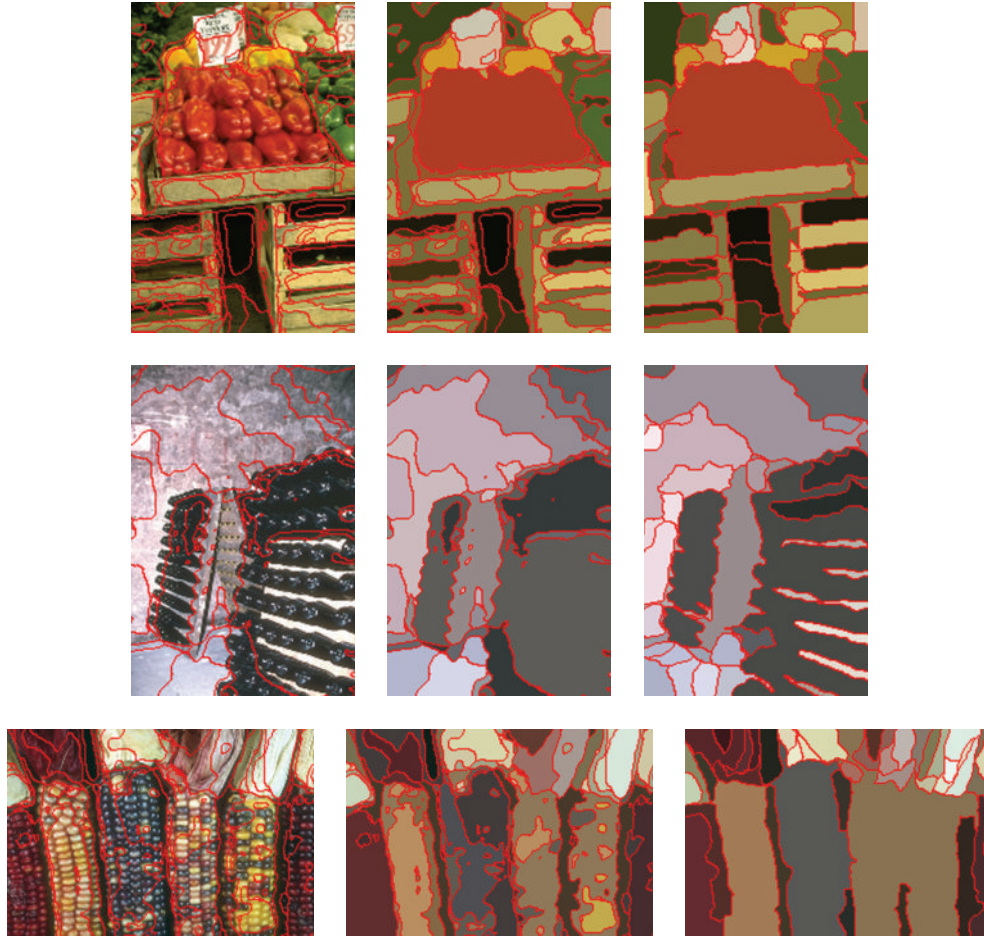


Figure 6: Objects images: left column: original images with the boundaries obtained by using the proposed algorithm; middle and right columns: the segmentation results obtained by using the proposed algorithm and the CTM algorithm, respectively.

3.2. The Modified CTM Algorithm

In this section, we modify the CTM algorithm [32] to segment the MQ feature vectors of an image. With a set of MQ feature vectors, the number of coding bits can be approximated as

$$L(\varepsilon) = \frac{N+D}{2} \log_2 \left(\det \left(I + \frac{D}{\varepsilon^2 N} \Sigma \right) \right) + \frac{D}{2} \log_2 \left(1 + \frac{\mu^T \mu}{\varepsilon^2} \right), \quad (3.4)$$

where μ is the mean vector, Σ is the covariance matrix, ε is the distortion incurred, D is the feature dimension, and N is the number of feature vectors. For K sets of MQ feature vectors, the total number of coding bits is given by

$$L_{\text{tot}}(\varepsilon) = \sum_{i=1}^K L_i(\varepsilon) - N_i \log_2 \left(\frac{N_i}{N} \right), \quad (3.5)$$

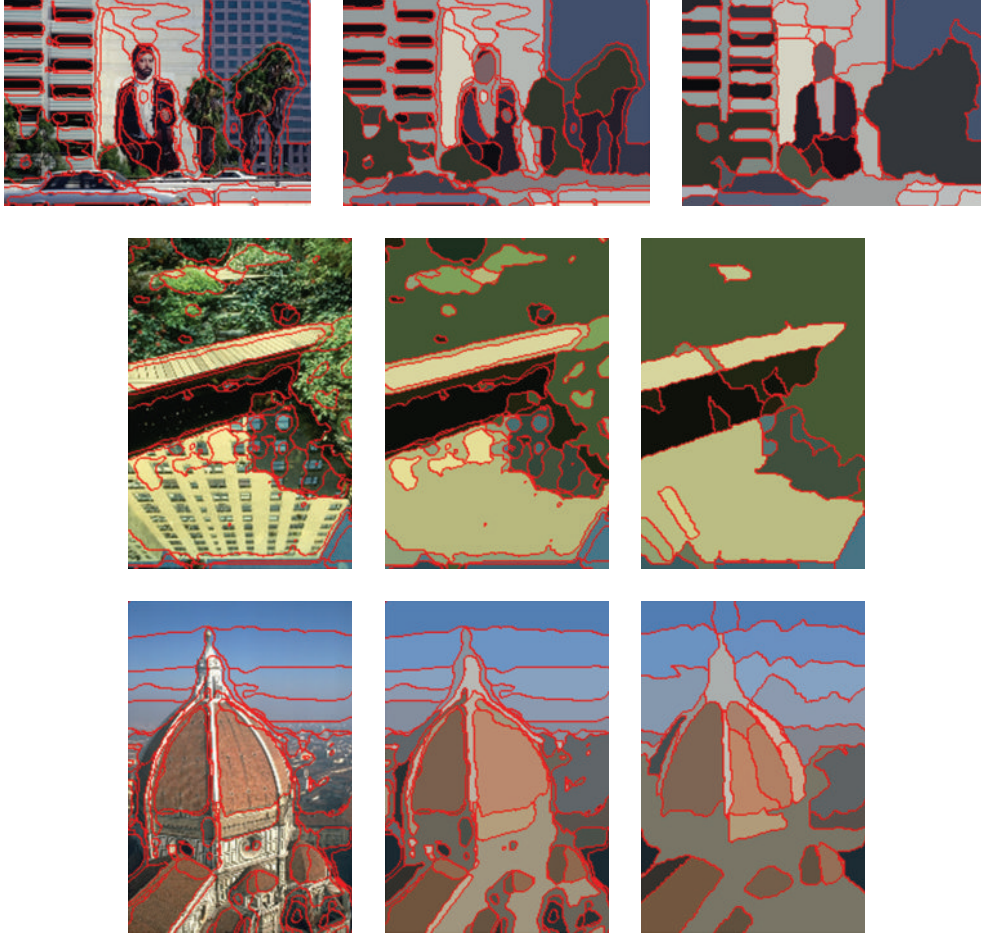


Figure 7: Urban images: left column: original images with the boundaries obtained by using the proposed algorithm; middle and right columns: the segmentation results obtained by using the proposed algorithm and the CTM algorithm, respectively.

where $L_i(\varepsilon)$ and N_i are the number of coding bits obtained by (3.4) and the number of MQ feature vectors in the i th set, respectively, and N is the total number of MQ feature vectors, that is, $N = \sum_{i=1}^K N_i$. The idea behind CTM is to merge two sets of feature vectors such that the coding bits can be reduced maximally. The pairwise merging procedure of CTM is performed iteratively until no merge can reduce the coding bits any more. As mentioned in [32], the termination of CTM is dependent on the distortion parameter, ε , which can be determined by

$$\varepsilon^* = \min\{\varepsilon : d(\varepsilon) \geq \gamma\}, \quad (3.6)$$

where $d(\varepsilon)$ is the distance between a pair of segments with respect to ε .

Motivated by the rate distortion theory, which has been widely used in embedded image coding for optimal rate control, we propose a simple scheme to determine the candidates of ε . Specifically, for a sequence of increasing distortion values: $\varepsilon_1 < \varepsilon_2 < \dots$, the

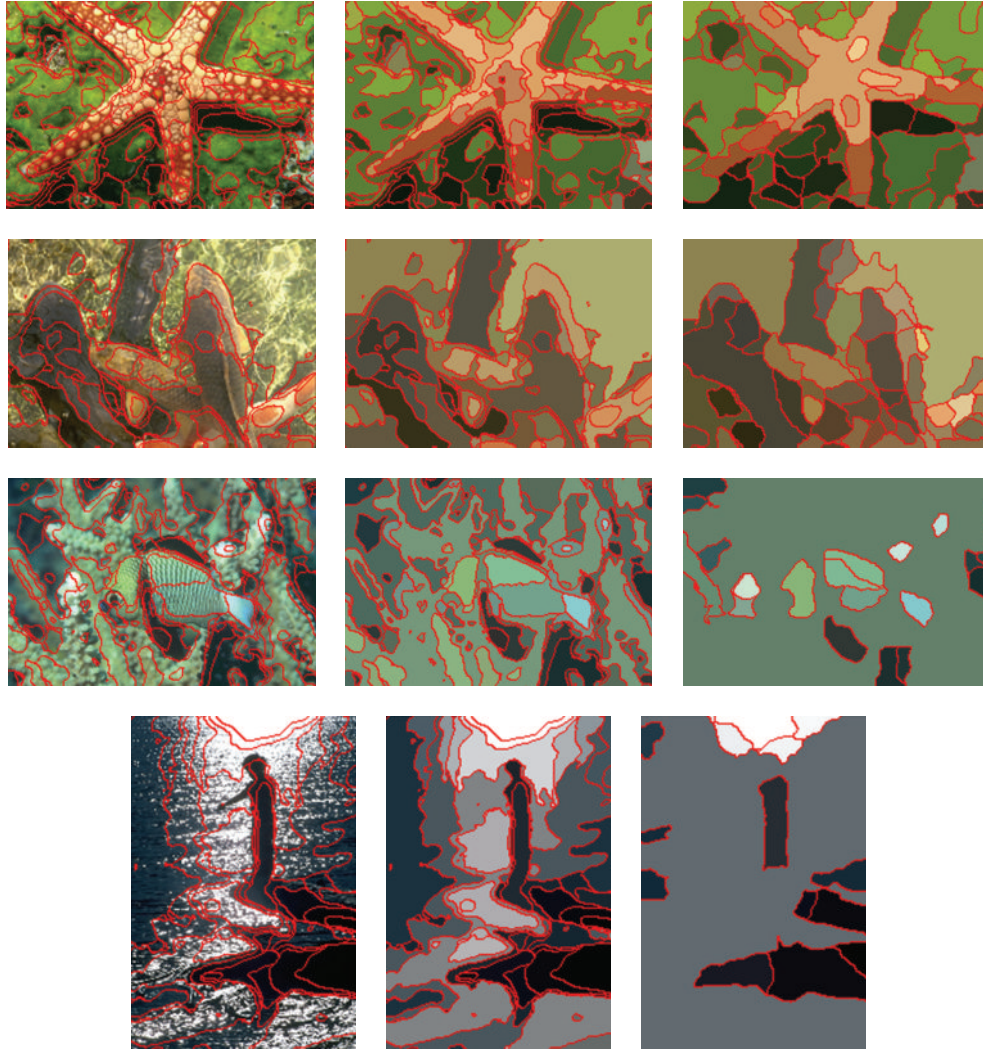


Figure 8: Water images: left column: original images with the boundaries obtained by using the proposed algorithm; middle and right columns: the segmentation results obtained by using the proposed algorithm and the CTM algorithm, respectively.

number of segments and the total number of coding bits are monotonically decreasing, that is, $K_1 > K_2 > \dots$ and $L_{\text{tot}}(\varepsilon_1) > L_{\text{tot}}(\varepsilon_2) > \dots$. The rate distortion slope (RDS) is thus defined as

$$S(\varepsilon_i) = \frac{\Delta D_i}{\Delta R_i}, \quad (3.7)$$

where

$$\Delta R_i = \frac{L_{\text{tot}}(\varepsilon_{i-1}) - L_{\text{tot}}(\varepsilon_i)}{N}, \quad (3.8)$$

$$\Delta D_i = \varepsilon_i - \varepsilon_{i-1},$$

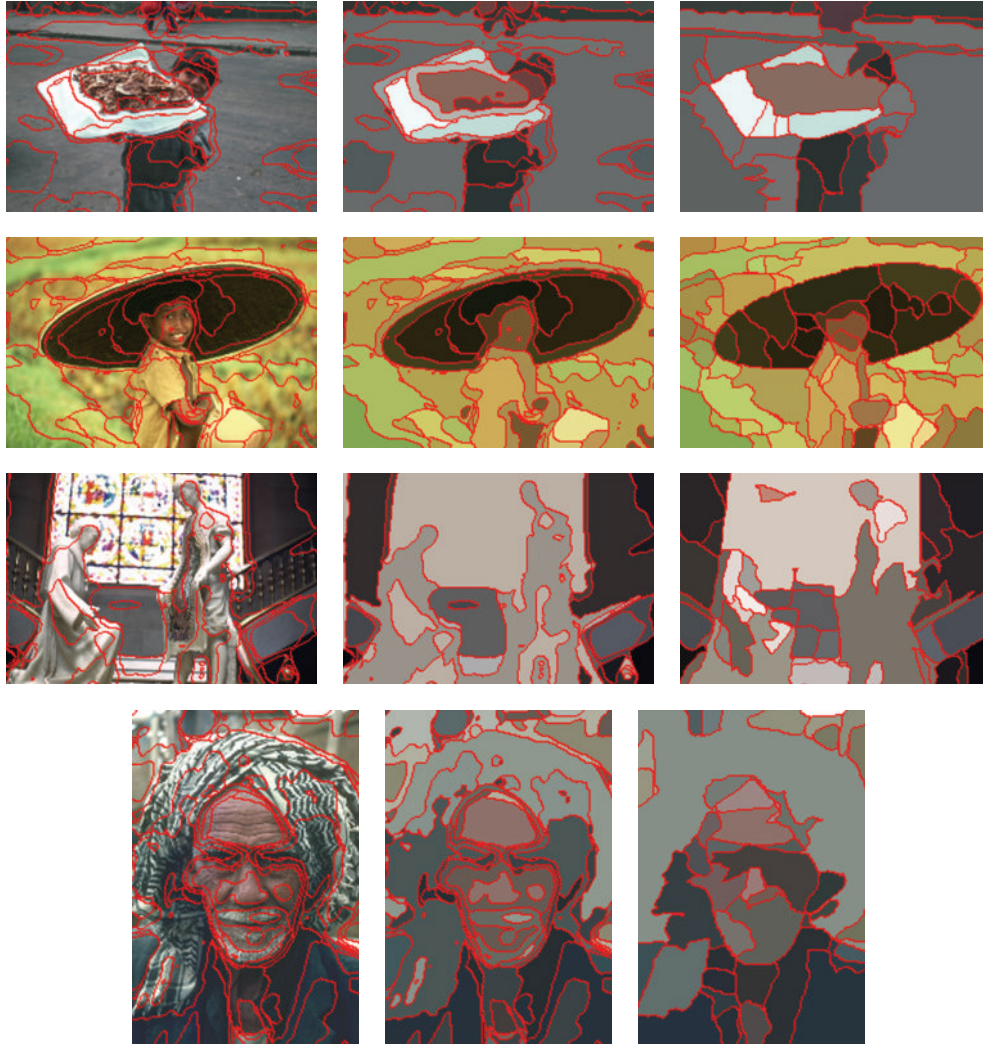


Figure 9: Portraits images: left column: original images with the boundaries obtained by using the proposed algorithm; middle and right columns: the segmentation results obtained by using the proposed algorithm and the CTM algorithm, respectively.

and N is the number of MQ feature vectors. As RDS should be nondecreasing, that is,

$$S(\varepsilon_i) \leq S(\varepsilon_{i+1}) \quad \text{for } \varepsilon_i < \varepsilon_{i+1}, \quad (3.9)$$

if $S(\varepsilon_i) > S(\varepsilon_{i+1})$, ε_i can be considered as a candidate to terminate the merging process of CTM. Thus, we modify the selection of ε^* as follows:

$$\varepsilon^* = \max_i \{ \varepsilon_i : S(\varepsilon_i) > S(\varepsilon_{i+1}), d(\varepsilon_i) \leq \gamma \}. \quad (3.10)$$

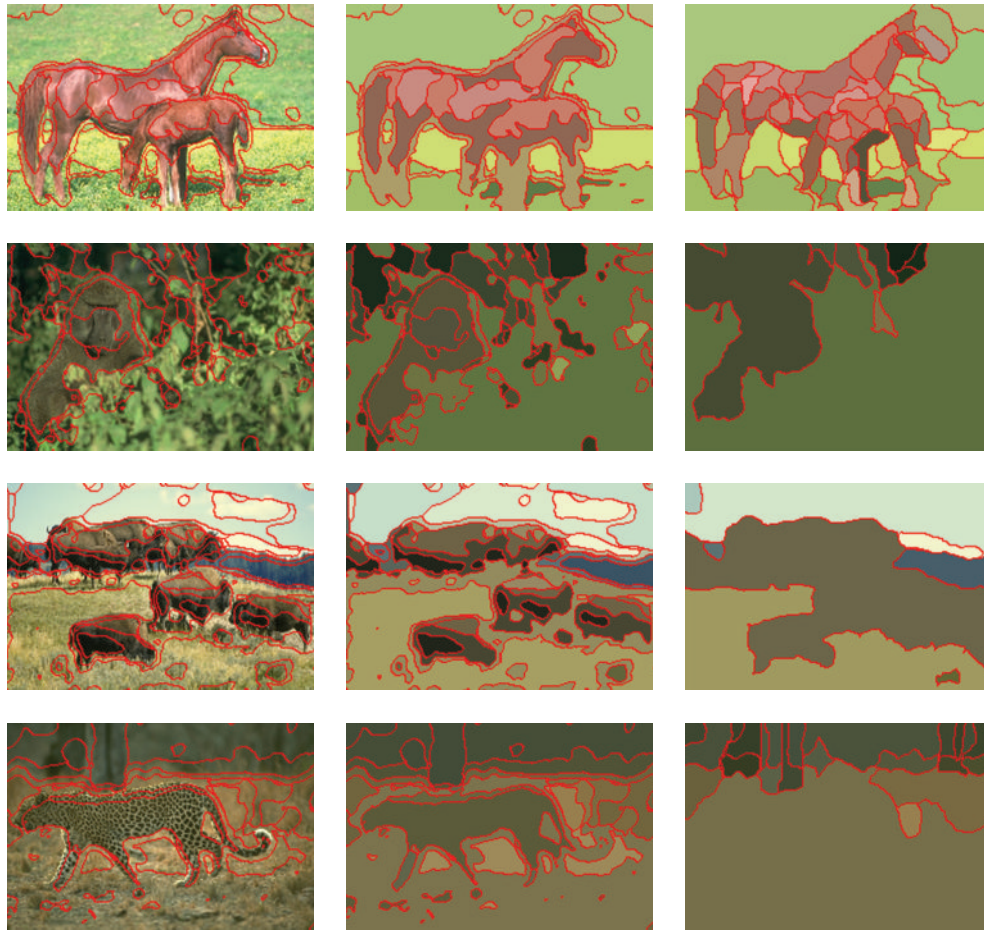


Figure 10: Animals images: left column: original images with the boundaries obtained by using the proposed algorithm; middle and right columns: the segmentation results obtained by using the proposed algorithm and the CTM algorithm, respectively.

Figure 1 depicts flowchart of the modified CTM with the RDS-based adaptive selection of ε^* , where the MQ coder-based image features are projected into a low-dimensional space via principal component analysis (PCA) in order to reduce the computational cost further, and the initial superpixels are obtained by using the simple K -means algorithm. Take the image shown in Figure 2(a) as an example; the candidates of ε are shown in Figure 2(e), where the horizontal and vertical axes are the distortion and the RDS values, respectively. Figures 2(b)–2(d) show the segmentation results with the first, second, and third candidates of ε . As one can see, the rate distortion information can be used to avoid overmerging of CTM.

4. Experimental Results

The proposed algorithm has been extensively evaluated on the Berkeley database [33]. The 9/7-wavelet filters adopted by JPEG2000 are used to extract the MQ coder-based image

Table 1: Average PRI and BDE on the Berkeley image database.

Measure	The proposed algorithm	CTM	Mean-shift	NCuts
Average PRI	0.757	0.756	0.755	0.723
Average BDE	9.2	9.4	9.7	9.6

features. The number of initial superpixels is set to 50. In addition to visual inspection, the boundary displacement error (BDE) and the probabilistic Rand index (PRI) [34] are used for quantitative evaluation. The segmentation results are compared with CTM, Mean-Shift and NCuts. In Mean-Shift, the parameters h_s and h_r are set to 13 and 19, respectively; in NCuts, the number of segments is 20. The threshold γ of CTM is set to 0.1, as suggested in [32].

We first evaluate the segmentation performance at various compression rates. Figure 3(a) shows a test image with two Brodatz textures, namely, wood and grass. Figure 3(b) depicts percentages of errors at various rates of bits per pixel (bpp). It is noted that the segmentation results even at low bpp rates are satisfactory; thus, a small portion of code stream is sufficient for the segmentation task. It has the advantage of saving transmission time, computational cost, and memory space, which are desirable especially for the Internet applications.

Table 1 shows the BDE and PRI performances compared to CTM, Mean-Shift, and NCuts. The proposed algorithm is preferable in terms of the average BDE.

The improvements in PRI and BDE using (3.10) are shown in Figures 4(a) and 4(b), respectively, where the horizontal axis is the threshold: γ . It is shown that the proposed algorithm is more robust by taking account of the rate distortion information to avoid overmerging.

Figures 5, 6, 7, 8, 9, and 10 are representative of the Landscape, Objects, Urban, Water, Portraits, and Animals images in the Berkeley database. The original images are shown in the left column. The segmentation results by using the proposed algorithm and the CTM algorithm are given in the middle and right columns, respectively. It is noted that, for images with high-detail contents, the proposed algorithm improves the segmentation results visually.

5. Conclusion

The MQ coder provides effective probability models, which is available at both encoder and decoder and therefore can be used to extract image features in the JPEG2000 domain directly. As a result, no overhead transmission is necessary to extract the feature vectors, and moreover the burden of decompressing a JPEG2000 image can be avoided. Based on the MQ coder, an efficient scheme of segmenting an image has been proposed. In order to avoid overmerging, the CTM algorithm has been modified by taking account of the rate distortion information. The proposed algorithm has been evaluated on images with Brodatz textures and the Berkeley image database. It is shown that the segmentation results at low-middle bpp rates are rather promising. In addition, for images with high-detail contents, the proposed algorithm is preferable in terms of the average BDE measure and visual comparison.

Acknowledgments

The authors are grateful to the maintainers of the Berkeley image database. The National Science Council of Taiwan under Grants NSC100-2628-E-239-002-MY2 supported this work.

References

- [1] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.
- [2] N. A. M. Isa, S. A. Salamah, and U. K. Ngah, "Adaptive fuzzy moving K-means clustering algorithm for image segmentation," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 4, pp. 2145–2153, 2009.
- [3] S. K. Mitra and J. F. Kaiser, *Handbook for Digital Signal Processing*, John Wiley & Sons, New York, NY, USA, 1993.
- [4] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedure*, John Wiley & Sons, New York, NY, USA, 3rd edition, 2000.
- [5] M. Li, W. S. Chen, and L. Han, "Correlation matching method for the weak stationarity test of LRD traffic," *Telecommunication Systems*, vol. 43, no. 3-4, pp. 181–195, 2010.
- [6] M. Li and W. Zhao, "Variance bound of ACF estimation of one block of fGn with LRD," *Mathematical Problems in Engineering*, vol. 2010, Article ID 560429, 14 pages, 2010.
- [7] M. Li, "A method for requiring block size for spectrum measurement of ocean surface waves," *IEEE Transactions on Instrumentation and Measurement*, vol. 55, no. 6, pp. 2207–2215, 2006.
- [8] T. N. Tan and A. G. Constantinides, "Texture analysis based on a human visual model," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '90)*, vol. 4, pp. 2137–2140, Albuquerque, NM, USA, April 1990.
- [9] D. E. Ilea and P. F. Whelan, "CTex—an adaptive unsupervised segmentation algorithm based on color-texture coherence," *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1926–1939, 2008.
- [10] S. Mallat, *A Wavelet Tour of Signal Processing*, Elsevier/Academic Press, Amsterdam, The Netherlands, 3rd edition, 2009.
- [11] M. K. Bashar, N. Ohnishi, and K. Agusa, "A new texture representation approach based on local feature saliency," *Pattern Recognition and Image Analysis*, vol. 17, no. 1, pp. 11–24, 2007.
- [12] C. M. Pun and M. C. Lee, "Extraction of shift invariant wavelet features for classification of images with different sizes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1228–1233, 2004.
- [13] J. Sun, D. Gu, S. Zhang, and Y. Chen, "Hidden Markov Bayesian texture segmentation using complex wavelet transform," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 151, no. 3, pp. 215–223, 2004.
- [14] 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 Transactions on Image Processing*, vol. 15, no. 10, pp. 3078–3088, 2006.
- [15] M. Li, W. Zhao, and S. Chen, "MBm-based scalings of traffic propagated in internet," *Mathematical Problems in Engineering*, vol. 2011, Article ID 389803, 21 pages, 2011.
- [16] M. Li and W. Zhao, "Representation of a stochastic traffic bound," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 9, pp. 1368–1372, 2010.
- [17] M. Li, "Generation of teletraffic of generalized Cauchy type," *Physica Scripta*, vol. 81, no. 2, Article ID 025007, 2010.
- [18] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, UK, 2nd edition, 2004.
- [19] T. Schreiber, "Interdisciplinary application of nonlinear time series methods," *Physics Report*, vol. 308, no. 1, pp. 1–64, 1999.
- [20] B. J. West, E. L. Geneston, and P. Grigolini, "Maximizing information exchange between complex networks," *Physics Reports*, vol. 468, no. 1–3, pp. 1–99, 2008.
- [21] M. Li, "Fractal time series—a tutorial review," *Mathematical Problems in Engineering*, vol. 2010, Article ID 157264, 26 pages, 2010.
- [22] S. Y. Chen, H. Tong, and C. Cattani, "Markov models for image labeling," *Mathematical Problems in Engineering*, vol. 2011, Article ID 814356, 18 pages, 2011.
- [23] S. Y. Chen, H. Tong, Z. Wang, S. Liu, M. Li, and B. Zhang, "Improved generalized belief propagation for vision processing," *Mathematical Problems in Engineering*, vol. 2011, Article ID 416963, 12 pages, 2011.
- [24] S. Y. Chen, J. Zhang, Q. Guan, and S. Liu, "Detection and amendment of shape distortions based on moment invariants for active shape models," *IET Image Processing*, vol. 5, no. 3, pp. 273–285, 2011.
- [25] F. R. Alexander, P. V. Juan, and S. Hichem, "Estimating quality bounds of JPEG 2000 compressed leukocytes images," in *Proceedings of the 2nd Mexican Conference on Pattern Recognition*, pp. 27–29, The National Institute of Astrophysics, Puebla, Mexico, 2010.

- [26] T. Y. Sung and H. C. Hsin, "A hybrid image coder based on SPIHT algorithm with embedded block coding," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E90-A, no. 12, pp. 2979–2984, 2007.
- [27] F. Zargari, A. Mosleh, and M. Ghanbari, "A fast and efficient compressed domain JPEG2000 image retrieval method," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 4, pp. 1886–1893, 2008.
- [28] H.-C. Hsin, "Texture segmentation in the joint photographic expert group 2000 domain," *IET Image Processing*, vol. 5, no. 6, pp. 554–559, 2011.
- [29] T. Acharya and P. S. Tsai, *JPEG2000 Standard for Image Compression*, John Wiley & Sons, New York, NY, USA, 2005.
- [30] F. Auli-Llinas, J. Serra-Sagristà, and J. Bartrina-Rapesta, "Enhanced JPEG2000 quality scalability through block-wise layer truncation," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, Article ID 803542, 11 pages, 2010.
- [31] A. K. Gupta, S. Nooshabadi, and D. Taubman, "Novel distortion estimation technique for hardware-based JPEG2000 encoder system," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 7, pp. 918–923, 2007.
- [32] A. Y. Yang, J. Wright, Y. Ma, and S. S. Sastry, "Unsupervised segmentation of natural images via lossy data compression," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 212–225, 2008.
- [33] http://www.eecs.berkeley.edu/~yang/software/lossy_segmentation/.
- [34] R. Unnikrishnan, C. Pantofaru, and M. Hebert, "Toward objective evaluation of image segmentation algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 929–944, 2007.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

