A FAST MQ TABLE BASED MERGING ALGORITHM FOR IMAGE SEGMENTATION

Hsi-Chin Hsin[1], *Tze-Yun Sung*[2] *and Yaw-Shih Shieh*[2]

Department of Computer Science and Information Engineering, National United University, Taiwan[1]
Department of Electronics Engineering, Chung Hua University, Taiwan[2]

ABSTRACT

This paper presents a simple scheme to segment an image in the compressed domain. Thus the burden of decompressing computation can be avoided. In addition, there is no need to transmit the segmentation result from encoder to decoder because the proposed method is based on the MQ table of JPEG2000, which is available at both encoder and decoder. Experimental results show that the running time can be reduced significantly and the boundary displacement error (BDE) measure can be improved.

Index Terms — Image segmentation, wavelet, rate-distortion slope, MQ coder

1. INTRODUCTION

Among numerous image segmentation algorithms, the feature based approach has received a lot of attention due largely to its computational efficiency [1]. Early research work on feature extraction is mainly at a single scale. It is however noted that an image is decomposed into band-pass sub-images by simple visual cortical cells in the human visual system (HVS) [2], which can be modeled by Gabor filters with spatial frequencies and orientations properly tuned [3]. Wavelet transform (WT) provides an efficient multiscale representation, which matches the characteristics of HVS [4]. Various WT-based algorithms were proposed to extract image features at multiple scales [5]-[6].

With the rapid growth of multimedia technologies and the Internet applications, image compression is still in great demand. It is thus desirable to extract image features in the compressed domain directly. The Joint Photographic Expert Group (JPEG) 2000 standard adopts WT as the underlying transform algorithm [7]. Pi proposed a simple scheme to estimate the global probability mass function (PMF) of wavelet subbands for image retrieval [8]. However, local PMF is needed for image segmentation. In [9], we proposed a simple method to estimate the local PMF of wavelet subbands, which can be used as image features.

Motivated by the idea behind the post compression rate distortion (PCRD) algorithm of JPEG2000, we propose a simple algorithm called the rate distortion based merging (RDM) algorithm for image segmentation. It can be applied to a JPEG2000 code stream instead of the decoded image. As a result, the burden of decompressing computation can be avoided. The remainder of the paper proceeds as follows. In Section 2, the JPEG2000 standard is reviewed briefly. In Section 3, the MQ table based rate distortion slope (MQRDS) is proposed to develop the RDM algorithm. Experimental results on the Berkeley color image database are given in Section 4. Conclusions can be found in Section 5.

2. REVIEW OF JPEG2000

The core of JPEG2000 is the embedded block coding with optimized truncation (EBCOT) algorithm [7], which adopts wavelet transform (WT) as the underlying method for subband decompositions. 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. Three coding passes, namely the significance propagation (SP) pass, the magnitude refinement (MR) pass and the clean up (CU) pass are performed in BPC. 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 neighboring coefficients, the MQ coder defines context labels with their respective probability modes stored in the MQ table.

In JPEG2000, a large image can be partitioned into non-overlapped sub-images 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 post compression rate distortion (PCRD) algorithm. Specifically, let $\{B_i\}$ be the code blocks of an image. The code stream of B_i can be terminated at some point, says 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: $R = \sum_i R_i^{n_i} \le R_c$, 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. Note that the coding passes with non-increasing rate distortion slopes (RDS) are candidates for the optimal truncation points. Motivated by the above, we propose an efficient image segmentation scheme in the following section.

3. THE PROPOSED ALGORITHM

As the binary wavelet variables of an image are almost independent across bit-planes, the probability mass function (PMF) [8] can be approximated by $P(|c| = x) = \prod_{j=0}^{n-1} P_j(x_j)$, where $x = \sum_{j=0}^{n-1} x_j \cdot 2^j; x_j \in \{0,1\}$ is the magnitude of a wavelet coefficient, c, $P_j(\circ)$ is the PMF of the binary wavelet variable, x_j, on the j^{th} bit-plane, and n is the number of bit-planes. For image segmentation, we proposed a simple method to estimate the local PMF based on the MQ table [9]. Specifically, the probability of 1-bit, $P_j(x_j = 1)$, is given by

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

where Qe_Value is the probability of the less probable symbol (LPS), which is stored in the MQ table, and MPS denotes the more probable symbol.

Motivated by the PCRD algorithm, wavelet coefficients with similar characteristics should be coded together and significant ones should be coded as soon as possible. It seems essential to take account of the rate distortion slope of wavelet segments for the image segmentation task. Thus we propose the MQ table based rate distortion slope (MQRDS) as follows.

$$S_m = \frac{E[D_m]}{E[L_m]} \tag{2}$$

$$D_m = \sum_{i=1}^{N_m} x_{m,i}^2 \tag{3}$$

$$x_{m,i} = \sum_{j=0}^{n-1} x_{m,i,j} \cdot 2^j \; ; x_{m,i,j} \in \{0,1\} \tag{4}$$

$$\begin{aligned} E[D_m] &= \sum_{i=1}^{N_m} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} E\left[x_{m,i,j} \cdot x_{m,i,k}\right] \cdot 2^{j+k} \\ &\cong \sum_{i=1}^{N_m} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} E\left[x_{m,i,j}\right] \cdot E\left[x_{m,i,k}\right] \cdot 2^{j+k} \end{aligned} \tag{5}$$

$$E[x_{m,i,j}] = P_{m,i,j}(x_{m,i,j} = 1) \tag{6}$$

$$E\left[L_m\right] = (D + N_m) \cdot E\left[R_m\right] - N_m \log_2\left(\frac{N_m}{N}\right) \tag{7}$$

$$E[R_m] = \sum_{j=0}^{n-1} H(x_{m,j}) \tag{8}$$

$$\begin{aligned} H(x_{m,j}) = &-P_{m,j}(x_{m,j} = 1) \cdot \log_2(P_{m,j}(x_{m,j} = 1)) \\ &- P_{m,j}(x_{m,j} = 0) \cdot \log_2(P_{m,j}(x_{m,j} = 0)) \end{aligned} \tag{9}$$

$$P_{m,j}(x_{m,j}) = \frac{1}{N_m} \sum_{i=1}^{N_m} \cdot P_{m,i,j}(x_{m,i,j}) \tag{10}$$

where m and i are indices for wavelet segments and wavelet pixels, respectively, j and k are indices for bit-planes, $x_{m,i}$ is the magnitude of wavelet pixel i within segment m, $x_{m,i,j}$ and $x_{m,i,k}$ are the binary variables of $x_{m,i}$ on bit-planes j and k; their respective probabilities $P_{m,i,j}(x_{m,i,j})$ and $P_{m,i,k}(x_{m,i,k})$ are almost independent across bit-planes, and can be obtained from the MQ table, n is the number of bit-planes, D is the feature space dimension, N_m is the number of wavelet pixels within segment m, $N = \sum_{m=1}^{K} N_m$ is the total number of wavelet pixels, $H(\circ)$ is the entropy operation, and equation (7) is similar to that in [10]. After merging two wavelet segments, says m and n, the change of MQRDS is given by

$$\Delta S_{mn} = \left[S_{mn} - \left(\frac{N_m}{N_m + N_n} S_m + \frac{N_n}{N_m + N_n} S_n \right) \right] \Big/ S_{mn} \tag{11}$$

where S_m and S_n are the MQRDS of wavelet segments, m and n, with sizes, N_m and N_n, respectively, and S_{mn} is the MQRDS of the merged wavelet segment. As one can see, the change of MQRDS is likely to be increased significantly for wavelet segments with similar characteristics. Thus, we propose a simple algorithm called the rate distortion based merging (RDM) algorithm for image segmentation, which is presented in steps below.

<u>The RDM algorithm</u>

Step 1: Given a JPEG2000 code stream, compute the MQ table based local PMF of wavelet coefficients using equation (1).

Step 2: As over-segmented regions known as superpixels are in general needed for any merging algorithms [10], the low-level initial segmentation that can be obtained by clustering the local PMF as features is thus performed.

Step 3: For all pairs of superpixels, compute their respective changes of MQRDS using equation (11), and merge the one with maximum change of MQRDS.

Step 4: Continue the merging process in step 3 until the change of MQRDS is insignificant.

In order to reduce the computation time, the following equation can be used to approximate equation (5).

$$\begin{aligned} E[D_m] \cong N_m \cdot [&\sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \left(\frac{1}{N_m} \sum_{i=1}^{N_m} P_{m,i,j}(x_{m,i,j} = 1) \right) \\ &\cdot \left(\frac{1}{N_m} \sum_{i=1}^{N_m} P_{m,i,k}(x_{m,i,k} = 1) \right) \cdot 2^{j+k}] \end{aligned} \tag{12}$$

Moreover, the cross terms of the above equation are not significant and therefore can be discarded for computational simplicity. Note that the MQ table defined in JPEG2000 is finite. Thus equation (9) can be obtained by look-up-table (LUT); this sure reduces the computation time further.

4. EXPERIMENTAL RESULTS

The RDM algorithm has been extensively evaluated on the Berkeley image database [11]. We adopted the Waveseg algorithm [12] to compute the initial supperpixels. In order to avoid decoding a JPEG2000 code stream, the Waveseg algorithm was applied to the estimated wavelet coefficients instead of the decoded wavelet coefficients. Specifically, the estimated wavelet coefficient of x_i using the MQ table-based local PMF is as follows.

$$E[x_i]=\sum_{j=0}^{n-1}E[x_{i,j}]\cdot 2^j=\sum_{j=0}^{n-1}P_{i,j}(x_{i,j}=1)\cdot 2^j \tag{13}$$

where $P_{i,j}(x_{i,j}=1)$ is the probability of 1-bit on the j^{th} bit-plane, which can be obtained from the MQ table. The resulting superpixels were merged by RDM with threshold, T_d, set to 0.1. We compared the RDM algorithm with two other state-of-the-art algorithms known as Mean-shift [13] and CTM [10]. In Mean-shift, the parameters, h_s and h_r, were set to 13 and 19, respectively; in CTM, the threshold γ was set to 0.1, as suggested in [10]. The original images shown at the top of Figure 1 are natural images contained in the Berkeley database, namely Landscape and Horses. Their respective segmentation results using RDM, CTM, and Mean-shift are shown in the second, third, and fourth row. Visual inspection shows that RDM and Mean-shift have similar performances.

In addition to visual inspection, the commonly used measure known as boundary displacement error (BDE) [14] was adopted for quantitative comparisons. BDE measures the average displacement error of boundaries between segmented images, which is nonnegative, and lower is better. It is noted that RDM outperforms CTM and Mean-shift in terms of the average BDE as shown in Table 1.

The running times on a PC are given in Table 2. It shows that RDM is faster than CTM and Mean-shift due largely to the simple computations of equations (7) and (12). Moreover, RDM can be applied to a JPEG2000 code stream directly while most algorithms such as Mean-shift and CTM are primarily applied to the original or decoded image; and it takes more time to decode a compressed image.

Table 1: Average BDE on the Berkeley database

RDM	CTM	Mean-shift
8.7	9.4	9.7

Table 2: Execution times

	Landscape	Horses
RDM	8.7 s	10.7 s
Mean-shift	27.5 s	20.7 s
CTM	17.2 s	57.6 s

5. CONCLUSION

The MQ table defined in the JPEG2000 standard provides useful information that can be used to compute the local probability mass function (PMF) of wavelet coefficients. A simple local PMF-based scheme has been proposed to estimate the rate distortion slope (RDS) of a wavelet segment. It is noted that the RDS is increased significantly after merging wavelet segments with similar characteristics into a single segment. In this paper, we propose the rate distortion based merging (RDM) algorithm to segment images in the framework of JPEG2000. RDM has been evaluated on the Berkeley image database. Experimental results show that RDM is preferable in terms of the average BDE measure. In addition, the running time of RDM, which includes the computation of superpixels and the merging process, is faster than Mean-shift and CTM. As RDM is based on the MQ table, which is available at both encoder and decoder, no overhead transmission is needed. RDM can be applied to a JPEG2000 code stream directly. Thus the burden of decompressing computation can be avoided, and memory space that is required to store the decompressed image is no longer necessary from the segmentation point of view.

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

6. REFERENCES

[1] A. K. Jain, "Data clustering: 50 years beyond K-means," Pattern Recognition Letters, vol. 31, pp. 651-666, 2010.

[2] T. N. Tan and A. G. Constantinides, "Texture analysis based on a human visual model," Proc. IEEE ICASSP, pp. 2091-2110, 1990.

[3] D. E. Ilea and P. F. Whelan, "CTex - An adaptive unsupervised segmentation algorithm based on color-texture coherence," IEEE Trans. IP, vol. 17, pp. 1926-1939, 2008.

[4] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, 2009.

[5] 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, pp. 11-24, 2007.

[6] 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.

[7] T. Acharya, and P. S. Tsai, JPEG2000 Standard for Image Compression, John Wiley & Sons, 2005.

[8] 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.

[9] H. C. Hsin, "Texture segmentation in the joint photographic expert group 2000 domain," IET Image Processing, vol. 5, pp. 554-559, 2011.

[10] A. Y. Yang, J. Wright, S. Sastry, and Y. Ma, "Unsupervised segmentation of natural images via lossy data compression," Computer Vision and Image Understanding, vol. 110, no. 2, pp. 212-225, 2008.

[11] www.eecs.berkeley.edu/~yang/software/lossy_segmentation/

[12] C. R. Jung, "Unsupervised multiscale segmentation of color images," Pattern Recognition Letters, 28, pp. 523-533, 2007.

[13] D. Comanicu and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Trans. PAMI, vol. 24, pp. 603-619, 2002.

[14] 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.

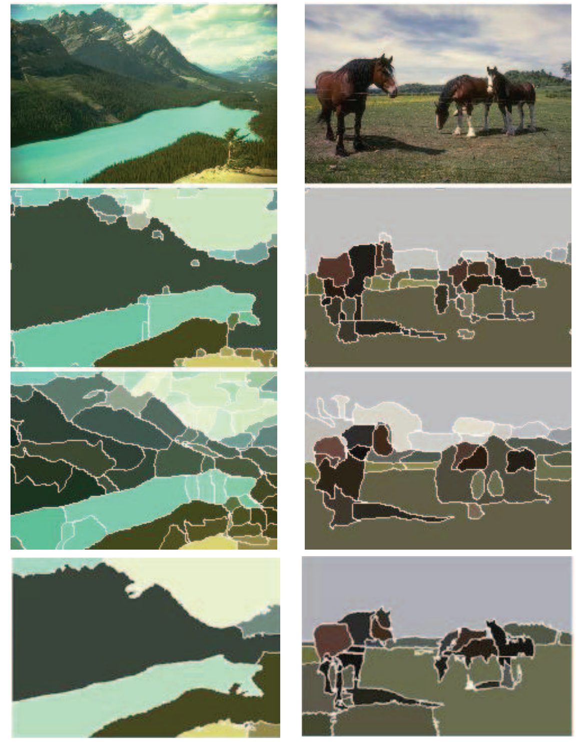

Figure 4: Top row: original images; Second row: segmentation using RDM; Third row: segmentation using CTM; Fourth row: segmentation using Mean-shift.