AN EMBEDDED SCALING-BASED ARBITRARY SHAPE REGION-OF-INTEREST CODING METHOD FOR JPEG2000

Mahesh M. Subedar, Lina J. Karam

Department of Electrical Engineering
Arizona State Univesity
Tempe, AZ 85287-7206

Glen P. Abousleman

Compression, Communications & Intelligence Lab
General Dynamics Decision Systems
Scottsdale, AZ 85257

ABSTRACT

The shape information of objects is perceptually very significant and can aid in the recognition of objects. Depending upon the bit budget, Internet browsing applications can take advantage of an arbitrary-shape region-of-interest (ROI) coding method by sending only the shape information first, and then progressively transmitting the ROI texture, followed by the less important non-ROI texture information. Unfortunately, the existing ROI coding methods, including the JPEG2000-based methods, do not support separate coding and decoding of the shape information. In particular, the maxshift and the scaling-based methods of JPEG2000 are limited in that the former does not allow the coding of any non-ROI bitplanes prior to the coding of the ROI bitplanes, and the latter supports only rectangular and circular ROI regions. This paper presents a novel ROI-based coding method that extends the scaling-based method of JPEG2000, and allows the coding of arbitrary-shape ROIs, as well as the separate coding and transmission of the ROI shape information prior to the texture information. Coding results are presented to illustrate the performance of the proposed ROI-based coding method.

1. INTRODUCTION

JPEG2000 [1, 2] supports two ROI coding methods [3, 4]—the maxshift method and the scaling-based method. The maxshift method has the advantage of not having to send the shape information to the decoder, but has the disadvantage of having to code all ROI regions prior to coding any part of the non-ROI region. If the bit budget is not sufficient, the non-ROI region is not coded at all. For many applications, such as medical imaging and automatic target detection, it is important to preserve the non-ROI information to some extent for contextual purposes. On the other hand, the scaling-based method has the advantage of allowing the partial coding of the non-ROI region prior to coding the entire ROI. In the present form, however, the scaling-based method of JPEG2000 supports only rectangular and circular ROIs, since the transmission of the entire binary mask to the decoder would be very expensive. Our proposed method overcomes this limitation by incorporating an efficient and less complex shape coding technique, which will encode the arbitrary shape contour information of the ROI regions, while consuming only a very small fraction of the overall bit budget.

Other ROI coding methods have been proposed in the literature. In [5] and [6], for example, ROI coding for the SPHIT and Quadtree coders are discussed, respectively. In [7, 8], a bitplane-by-bitplane shift method is proposed, which has the option of shifting the ROI bitplanes by a scale factor, followed by the alternating transmission of the non-ROI and ROI bitplanes. In [9], a similar method has been proposed, which differs from the method of [7, 8] in the way the ROI and non-ROI bitplanes are arranged. In [10], a new approach to the maxshift method is given, which shifts the ROI bitplanes by arbitrary values. The ROI bitplanes overlapping with the non-ROI bitplanes are replaced with zeros, which results in the loss of a significant amount of ROI information for smaller shift values.

This paper is organized as follows. Section 2 gives a brief overview of the ROI coding methods that are incorporated in the JPEG2000 standard. In Section 3, our proposed arbitrary-shape ROI coding method is described. Coding results are presented in Section 4, and a conclusion is given in Section 5.

2. OVERVIEW OF JPEG2000 ROI CODING METHODS

JPEG2000 [1] is a wavelet-based bitplane coding method. In this algorithm, the original image samples are shifted in level (if they are unsigned pixel values) such that they form a symmetric distribution of the discrete wavelet transform (DWT) coefficients for the LL sub-band. If lossy compression is chosen, the transformed coefficients are then quantized. The ROI bitplanes of the quantized coefficients are shifted up according to the applied ROI coding method. The bitplanes are coded from the most significant bitplane (MSB) to the least significant bitplane (LSB) using the embedded block coding with optimal truncation (EBCOT) algorithm. The EBCOT [11] algorithm has three coding passes—significance propagation pass, magnitude refinement pass, and the cleanup pass. The significance propagation pass codes the significance of each sample based upon the significance of the neighboring eight pixels. The sign coding primitive is applied to code the sign information when a sample is coded for the first time as a non-zero bitplane coefficient. The magnitude refinement pass codes only those pixels that have already become significant. The cleanup pass will code the remaining coefficients that are not coded during the first two passes. The output symbols from each pass are entropy coded using context-based arithmetic coding. After all bitplanes are coded, the post-compression rate-distortion optimization (PCRD-opt) algorithm is applied to determine the contribution of each coding block to the final bitstream.

As stated earlier, JPEG2000 supports two ROI coding methods, as shown in Fig. 1. The maxshift method, which is given in part-I of JPEG2000 [1], shifts the ROI bitplanes above the non-ROI (background) bitplanes. This method can support ROIs of arbitrary shape, since there is no need to send the shape information to the decoder. However, to ensure that there is no ambiguity between ROI and non-ROI bitplanes, all ROI bitplanes are shifted

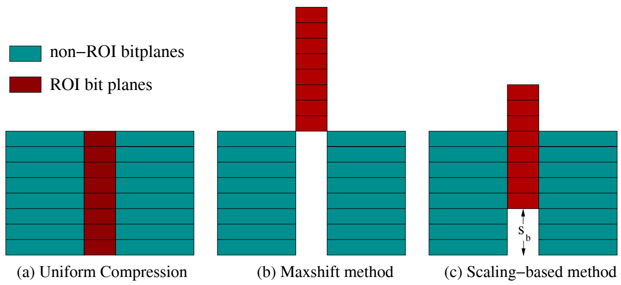

Fig. 1. (a) Uniform compression, (b) bitplane shifts for the maxshift method, and (c) bitplane shifts for the scaling-based method.

above the non-ROI bitplanes. Hence, the decoder needs to decode all the ROI regions before decoding any of the the non-ROI regions. Also, the decoder must support higher-precision arithmetic to decode all of the bitplanes. Note that although ROIs of arbitrary shape can be coded, this method does not support coding and decoding of the ROI shape information.

Part-II of the JPEG2000 [2] standard supports the scaling-based method with rectangular and circular ROI regions. At the encoder, the ROI bitplanes can be shifted up with an arbitrary scale factor, s_b, where s_b ranges from 1 to the maximum number of bitplanes. The shape information and the scale factor are sent to the decoder, where the ROI bitplanes are shifted down before decoding. For the scaling-based method, only rectangular and circular ROIs are supported, since the ROI mask overhead information required by the decoder would otherwise occupy a large portion of the bitstream.

In this paper, we address this inherent limitation of JPEG2000 by efficiently coding the boundary of each ROI. The mask (ROI shape) can then be reconstructed at the decoder from the ROI boundary using an efficient filling algorithm without having to include extra side information in the bitstream.

3. ARBITRARY-SHAPE ROI CODING METHOD FOR JPEG2000

The block diagram of the proposed coding method is shown in Fig. 2. The proposed method facilitates the coding of the arbitrary-shaped ROIs with any bitplane scaling factor, s_b. The shape information of each ROI is then embedded in the final bitstream. The encoding steps for the proposed method are as follows:

- Optional down/up sampling of the ROI mask for the efficient coding of large or multiple ROIs;
- Differential Chain Coding/Decoding of the ROI boundaries;
- Wavelet mask generation from the ROI boundaries;
- Bitplane shifting of the ROI regions using the desired scale factor, s_b;
- JPEG2000-based EBCOT coding.

3.1. Down/up sampling of the ROI mask

For very-low-bit-rate and/or multiple ROI applications, a down-sampling operation [12] can be applied to reduce the boundary lengths. The resulting ROI mask degradation depends upon the down-sampling factor. The down-sampling operation consists of reducing the size of the original bounding rectangle (minimum rectangle enclosing the arbitrarily-shaped ROI) from $m \times n$ to $m_d \times n_d$, where $0 < m_d < m$ and $0 < n_d < n$. The horizontal and vertical downsampling factors are given by $s_m \cong \frac{m_d}{m}$ and $s_n \cong \frac{n_d}{n}$, respectively, where s_m,$s_n$$\in$ [0,1). If down sampling is applied at the encoder, then the original integer ROI bounding box dimensions, m and n, need to be sent to the decoder to calculate s_m and s_n, so that the corresponding upsampling operation can be performed.

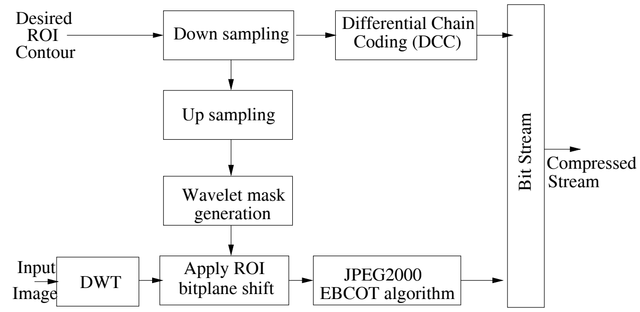

Fig. 2. Proposed embedded arbitrary-shape ROI coding method.

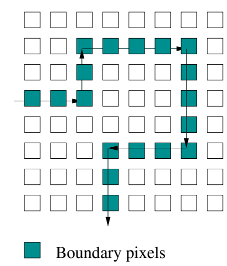

Fig. 3. Differential chain coding of ROI boundary.

3.2. Differential Chain Coding (DCC)

The differential chain coding (DCC) method of [13] is adopted to code the ROI contour information. Beginning from a seed point located at the top left-most contour pixel, the DCC algorithm traverses the entire contour until it reaches the starting point, as shown in Fig. 3. As compared to the previous movement, the direction to the next pixel is noted as either the same direction (SD) or a different direction (DD). The different direction may be either counter-clockwise (CCW) or clockwise (CW). Accordingly, the movements are Huffman coded as follows:

SD = 0; DDCW = 10 & DDCCW = 11.

DCC provides lossless compression of the contour information. Differential chain decoding is then applied at the decoder to reconstruct the ROI contours.

3.3. Wavelet mask generation from the ROI boundaries

Given the ROI contours, an efficient filling algorithm [13] is applied to obtain the spatial ROI mask of arbitrary shape. The corresponding wavelet-domain ROI mask is then generated from the spatial-domain ROI mask using a wavelet mask generation method similar to the one used in the maxshift method of JPEG2000. An example wavelet mask generated for a 2-level DWT decomposition is shown in Fig. 4.

(a) ROI mask (b) Wavelet mask

Fig. 4. (a) ROI mask, and (b) Wavelet mask for 2 levels of DWT.

3.4. Bitplane shifting of the ROI regions

Once the wavelet mask is generated for all sub-bands, the wavelet coefficients that are specified by the wavelet mask are scaled up by s_b with respect to the non-ROI region. The scaling factor can be chosen depending upon the bit budget for the ROI and non-ROI regions.

3.5. JPEG2000-based EBCOT coding

The scaled-up DWT coefficients of the input image are fed to the JPEG2000-based EBCOT coding algorithm [11] as shown in Fig. 2. In our present implementation, the Comment and Extension (CME) marker is used to send the shape information to the decoder.

4. RESULTS

To illustrate the performance of the proposed method, ROIs of arbitrary shape in different images were coded. Coding results are presented for the 2048×2560 "bike", and 512×512 "f16" images. The presented PSNR results are compared over the observed ROIs. Fig. 5 shows the reconstructed "bike" image, with the glass of wine being the ROI, at a bit rate of 0.008 bits/pixel (bpp), and an ROI scale factor $s_b = 8$. Note that the maximum bitplane index for the bike image is 14. It can be seen that the proposed method results in better reconstruction of edges and is thus of better perceptual quality as compared to the JPEG2000 scaling-based ROI coding method. In addition, the proposed method results in a higher PSNR over the desired ROI and is able to reconstruct more background information than the JPEG2000 scaling-based ROI coding method. Fig. 7 illustrates the progressive transmission and decoding of the "f16" image with the airplane being the ROI. At extremely low bit rates, the ROI mask is decoded first (Fig. 7(b)), with the ROI texture information being progressively decoded thereafter (Fig. 7(c)).

5. CONCLUSION

We have presented a new and novel image coding algorithm that extends the scaling-based method of JPEG2000, and supports the coding of arbitrarily-shaped regions of interest, as well as the separate coding of the ROI shape information. The proposed method allows the progressive transmission of the ROI shape information, followed by the texture information. Performance results illustrate that the proposed algorithm significantly outperforms the scaling-based ROI coding method of JPEG2000.

6. REFERENCES

[1] ISO/IEC 15444-1, "JPEG2000 image coding system – Part 1: core coding system," Tech. Rep., ISO, 2000.

[2] ISO/IEC JTC1/SC20 WG1 N2000, "JPEG2000 Part 2 final committee draft," Tech. Rep., ISO, 2000.

[3] C. Christopoulos, J. Askelof, and M. Larsson, "Efficient methods for encoding regions of interest in the upcoming JPEG2000 still image coding standard," *IEEE Signal Processing Letters*, vol. 9, no. 7, pp. 247–249, Sep 2000.

[4] C. Christopoulos, J. Askelof, and M. Larsson, "Efficient region of interest coding techniques in the upcoming JPEG2000 still image coding standard," in *IEEE International Conference on Image Processing*, 2000, vol. 2, pp. 41–44.

[5] E. Atsumi and N. Farvardin, "Lossy/lossless region-of-interest image coding based on set partitioning in hierarchical trees," in *IEEE International Conference on Image Processing*, Oct. 1998, vol. 1, pp. 87–91.

[6] D. Giguet, G.P. Abousleman, and L.J. Karam, "Very low bit-rate target-based image coding," in *Thirty-Fifth Asilomar Conference on Signals, Systems, and Computers*, Oct/Nov 2001.

[7] Z. Wang and A.C. Bovik, "Bitplane-by-bitplane shift (bbb-shift) - a suggestion for JPEG2000 region of interest image coding," *IEEE Signal Processing Letters*, vol. 9, no. 5, pp. 160–162, May 2002.

[8] Z. Wang, S. Banerfee, B.L. Evans, and A.C. Bovik, "Generalized bitplane-by-bitplane shift method for JPEG2000 ROI coding," in *IEEE International Conference on Image Processing*, Oct. 2002, vol. 3, pp. 81–84.

[9] L. Lijie and F. Guoliang, "A new method for JPEG2000 region-of-interest image coding: most significant bitplanes shift," in *Midwest Symposium on Circuits and Systems*, 2002, vol. 2, pp. 176–179.

[10] R. Grosbois, D. Santa-Cruz, and T. Ebrahimi, "New approach to JPEG2000 compliant region of interest coding," in *Proc. of SPIE, Applications of Digital Image Processing XXIV*, Aug. 2001, vol. 4472, pp. 267–275.

[11] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1151–1170, July 2000.

[12] David Giguet, "Error-resilient and very low bit rate image coding," *Master's Thesis, Arizona State University*, Dec 2000.

[13] Y.M.Y. Hasan and L.J. Karam, "Morphological reversible contour representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 3, pp. 227–240, Mar 2000.

(a) Original Image | (b) JPEG2000 scaling-based method, ROI PSNR = 29.97 dB | (c) Proposed method, ROI PSNR = 31.50 dB

Fig. 5. Reconstructed 2048×2560 "bike" images at a bit rate of 0.008 bpp and a scale factor s_b = 8, with the ROI being the glass of wine.

(a) Original | (b) JPEG2000 scaling-based method | (c) Proposed method

Fig. 6. Part of the reconstructed "bike" image of Fig. 5 showing part of the ROI region at a bit rate of 0.008 bpp and a scale factor s_b = 8.

(a) Original Image | (b) Reconstructed image at a bit rate of 0.0072 bpp | (c) Reconstructed image at a bit rate of 0.08 bpp and a scale factor s_b=5, ROI PSNR = 28.02 dB

Fig. 7. Coding results for the 512×512 "f16" image with the proposed method at different bit rates.