A NEW METHOD FOR JPEG2000 REGION-OF-INTEREST IMAGE CODING: MOST SIGNIFICANT BITPLANES SHIFT

Lijie Liu and Guoliang Fan
School of Electrical and Computer Engineering
Oklahoma State University, Stillwater, OK 74078
Email: {lijie, glfan}@okstate.edu

ABSTRACT

Region-of-interest (ROI) image coding is one of the new features included in the JPEG2000 image coding standard, where two methods are defined, i.e., the Maxshift method and the generic scaling based method. In this paper, we will propose a new JPEG2000 compliant ROI coding method called Most Significant Bitplane Shift (MSBShift) that combines the advantages of the two standard methods. The MSBShift method not only supports arbitrarily shaped ROI coding without coding the shape, but also enables the flexible adjustment of compression quality in the ROI and background. Additionally, the new method can efficiently code multiple ROIs with different priorities in an image.

1. INTRODUCTION

Region of Interest (ROI) image coding is one of new features included in the JPEG2000 standard, a new wavelet-based still image compression standard [1, 2]. This feature allows ROIs to be coded with better quality than the rest of the image, i.e., background (BG), implying a non-uniform quality distribution in a coded image. In JPEG2000, ROI coding is obtained by scaling down the BG coefficients in the wavelet domain so that ROI coefficients can be coded firstly in the following embedded bitplane coding [3, 4, 5]. Two scaling methods are defined in JPEG2000: the maximum shift (Maxshift) method [1] and the generic scaling based method [2].

In the generic scaling based method, the bits of BG coefficients are down shifted toward the least significant bit planes so that the ROI associated bits are placed in the higher bitplanes, as illustrated in Fig. 1(b). At the decoder, the bitplanes are firstly reconstructed from Most Significant Bitplane (MSB) to Least Significant Bitplane (LSB). Then according to the shape information of ROI included in the codestream, the BG coefficients are located and scaled up to their original places before the inverse wavelet transform is applied. As any scaling value is supported, the generic scaling based method allows the fine control on the relative importance between ROI and BG. However, the ROI shape information has to be coded, and this significantly increases the complexity and reduces the coding efficiency. In order to avoid this problem, the current standard restricts the ROI shape to be rectangle and ellipse [2, 6], and this restriction may limit the real application of ROI coding. On the other hand, in the generic scaling based method, the different wavelet subbands must have the same ROI definition, which is not desired in some applications [7]. The Maxshift method [1, 4] is the other standard method in JPEG2000. It may be considered as a particular case of the generic scaling based method when the scaling value is so large that there is no overlap between BG and ROI bitplanes, as seen from Fig. 1(c). It means that the scaling value, \( s \), must satisfy (1):

\[
s \geq \max(M_b),
\]

where \( M_b \) is the nominal maximum number of magnitude bitplanes in subband \( b \), and \( \max(M_b) \) is the largest magnitude for any coefficients [1]. After scaling, all bits of the ROI coefficients will be in higher bitplanes than all the bits associated with BG. At the decoder, the BG and ROI coefficients can be identified easily according to the bitplane positions, where all coefficients that are lower than the \( s^{th} \) bitplane are considered to belong to BG. Thus the shape of ROI is implicit for the decoder, and arbitrarily shaped ROI coding can be supported. Moreover, the Maxshift method can flexibly treat wavelet subbands differently, where different ROIs can be defined [5]. Compared with the generic scaling based method, the Maxshift method uses larger scaling values and reduces the compression efficiency by introducing more bitplanes [5, 6]. Unlike the generic scaling based method, the Maxshift method cannot control the relative importance between ROIs and the BG. No BG information can be received until all ROI coefficients are decoded fully. Moreover, the Maxshift method does not support multiple ROI coding, where multiple ROIs with different priorities may be involved. Considering the limitations of these two standard methods defined in JPEG2000, some authors proposed several improved methods for ROI image coding. A Maxshift-like method with low scaling values was proposed in [6]. This method removes all the overlap bitplanes between ROI and BG coefficients, and relatively modifies the quantization step size of the coefficients. The modification would bring the reduction of the final ROI quality. Wang et al. proposed a bitplane-by-bitplane shift (BbBShift) method in [7]. The BbBShift method shifts the bitplanes on a bitplane-by-bitplane basis instead of shifting them...
all at once in the Maxshift method. The BbBShift method supports arbitrarily shaped ROI coding without coding shape information, and it can finer control the ROI and BG quality than the Maxshift method. Unfortunately, it is not compatible with the current JPEG2000 ROI coding definitions and a new mode is needed. Moreover, the BbBShift method does not support multiple ROI coding, and its compression efficiency is similar to that of the Maxshift method [7].

In this paper, a new and flexible scaling based method called Most Significant Bitplane Shift (MSBShift) is proposed. Like the Maxshift method, the proposed MSBShift method can support arbitrary ROI shapes without coding shape information. The same as the generic scaling based method, the MSBShift method can flexibly adjust the relative importance between ROI and BG by using different scaling values. More interestingly, the new method supports multiple ROI coding.

2. MOST SIGNIFICANT BITPLANE SHIFT METHOD

The proposed MSBShift method is mainly based on the facts that at low bit rates, ROI in an image is desired to sustain higher quality than BG, while at the high bit rates, both ROI and BG can be coded with high quality and the difference between them is not very noticeable. So we can just isolate a certain number bitplanes of ROI bits in the most significant bitplanes to adjust the importance between ROI and BG. That means we only need to shift part of the most significant bitplanes of ROI coefficients instead of shifting the whole bitplanes of the ROI coefficients as the standard methods do. An illustration of the MSBShift method is shown in Fig. 2 (where $s = 6$). The whole bitplanes of ROI coefficients are divided into two parts: the most significant bitplanes and the left significant bitplanes. The number of the most significant bitplanes is represented as follows (2):

$$N_{th_{bs}} = \begin{cases} M_b - s & \text{if } s \leq M_b; \\ 0 & \text{if } s > M_b. \end{cases} \tag{2}$$

At the encoder, the most significant bitplanes of ROI coefficients are not shifted while the left significant bitplanes of ROI coefficients are down shifted towards the least significant bitplane with the BG coefficients. At the decoder, the ROI coefficients can be identified in the same way as the Maxshift method. All bits lower than the $s^{th}$ bitplane are shifted in the same way as the Maxshift method, and the result would be close to the Maxshift method. Therefore, the proposed MSBShift method may lead to an efficient and flexible ROI coding that can be tailored for different applications.

2.1. MSBShift ROI Coding

The proposed MSBShift method can code the ROI in one image with higher or the same quality as BG. As shown in Fig. 2, if the encoded bitstream is truncated or the encoding/decoding process is terminated before $2s$ most significant bitplanes are coded/decoded, the ROI will have higher quality than the BG. After $2s$ most significant bitplanes are coded/decoded (when the bitplanes of ROI and BG are overlap), the ROI and the BG will be coded with the same quality. The scaling value $s$ controls the relative quality between the ROI and BG. For example, we apply the different scaling values ($s = 6$ and $s = 5$) to the “Barbara” image ($720 \times 576, 8$bpp), where a rectangular ROI has been defined on the face region, Fig. 3 illustrates the ROI and BG distortion reduction with the increase of the decoding bit rate. When $s$ is small, e.g., $s = 5$, the quality of ROI and BG has no much difference (Fig. 3(b)). With the increase of the scaling value, e.g., $s > 6$ (Fig. 3(c) - (e)), ROI can be coded with higher quality than BG within a large range of bit rates. When all bitplanes are decoded, the lossless decoding is reached. If the scaling value $s \geq \max(M_b)$ (here $\max(M_b) = 12$), the result is the same as using the Maxshift method, as shown in Fig. 3(f). So the MSBShift can also generalize the Maxshift method in a way that is different from the generic scaling based method.

2.2. Functionality

Compared with the generic scaling based method, which only supports rectangle and ellipse ROI shapes [2], the proposed MSBShift method can code arbitrarily shaped ROI without coding the shape. The new method can also allow different ROI definitions in different wavelet subbands as the Maxshift method does. Different from the Maxshift method where no BG information can be received until all the ROI coefficients are reconstructed, the MSBShift method can flexibly adjust the relative importance between ROI and BG by using different scaling values. For example, Fig. 4 shows the “Barbara” image coded at 0.5bpp using the Maxshift method ($s = 12$) and the MSBShift method ($s = 10$). It can be observed that without visual difference at ROI, the MSBShift method provides better quality at BG than the Maxshift method. If ROI needs to have higher quality than BG, the large scaling value can be used and the results would be close to the Maxshift method. Therefore, the proposed MSBShift method may lead to an efficient and flexible ROI coding that can be tailored for different applications.

2.3. Multiple ROI Coding

Multiple ROI coding requires multiple ROIs to be coded with different quality according to their priorities in an image. The generic scaling based method can support multiple ROI coding. But it needs to code the shape and restrict the ROI shape. The Maxshift method cannot support multiple ROI coding, since all ROIs have to have the same scaling values in an image [7]. The proposed MSBShift method can support multiple ROI coding by scaling the different numbers of most significant bitplanes for different ROIs, as illustrated in Fig. 5. For example, three ROIs, i.e., ROI-1, ROI-2, and ROI-3, are defined in “Bike” ($2048 \times 2560, 8$bpp), shown in Fig. 6. The priority order of these ROIs is ROI-1>ROI-2>ROI-3. The upshifted numbers of the most significant bitplanes should be chosen as $s_1 > s_2 > s_3$, e.g., $s_1 = 10, s_2 = 8, s_3 = 6$. The scaling value stored in the codestream is $s = \max(s_1, s_2, s_3) = 10$. The reconstructed quality (PSNR) of three ROIs and BG is shown

![Diagram](image-url)
Fig. 3. The distortion reduction of ROI and BG with respect to the increase of bit rates at different scaling values, i.e., $s = 0, 5, 7, 9, 11, 13.$

in Fig. 7(a)(b). From Fig. 7, it can be found that at the low bit rates (e.g., bpp<0.9), all ROIs have the higher quality than BG. ROI-1 has the highest quality while ROI-3 has the lowest quality among three ROIs. When the bit rate increases, the BG quality increase quickly and is better than the quality of ROI-3 and even better than ROI-2. This is because the upshifted numbers of the most significant bitplanes of ROI-3 and ROI-2 are not large enough. At last, three ROIs and BG can reach the lossless quality at the same time when all the bitplanes are decoded. Hence, the MSBSHift method can support multiple ROI coding in a certain range of bit rates, which depends on the number of upshifted bitplanes for each ROI.

2.4. Complexity and Coding Efficiency

Since it is not necessary for the MSBSHift method to code ROI shape, the complexity is less than the generic scaling based method, and the coding efficiency is higher when the same scaling value is used. Compared with the Maxshift method, a little more complicated procedure need to be included in the MSBSHift method in order to shift back and to reconstruct the original bitplanes. If the point of lossless coding is reached, the bit rate produced by the MSBSHift method is not larger than the Maxshift method, because the MSBSHift method encodes less or at most the same number of bitplanes. The codestream generated by the proposed method is compliant with the current JPEG2000 format. It only needs a minor modification of the JPEG2000 Part I decoder as described before, and the modified decoder can also handle the codestream produced by the Maxshift method.

3. CONCLUSIONS

In this paper, we have proposed a so-called MSBSHift ROI coding method that has three primary advantages. Firstly, the MSBSHift method supports arbitrary ROI shapes without coding shape information. Secondly, it can control the relative importance between the ROIs and BG by using appropriate scaling values. Thirdly, the new method can code multiple ROIs with different priorities in an image at the low bit rates. We expect this idea is valuable for future research in ROI image coding and its applications.

4. REFERENCES

Fig. 4. “Barbara” coded at 0.5bpp using the Maxshift method (top, $s = 12$) and the MSBShift method (bottom, $s = 10$). The ROI is at the face region and 1/14 of the image size.

Fig. 5. MSBShift method for Multiple ROI coding.

Fig. 6. The “Bike” image with three ROIs.

Fig. 7. Multiple ROI coding results of “Bike” image at low bit rates (top) and mediate bit rates (bottom).