제25회 한국정보처리학회 춘계학술발표대회 논문집 제13권 제1호 (2006. 5)

A study of Region of Interest coding in JPEG2000

Yang Ou*, Ki-Jun Kang*, Bu-Kwon Lee*
* Dept. of Computer Science, Gyeongsang National University
e-mail : oy9798@hotmail.com

Abstract

One of the most significant characteristics of JPEG2000, the emerging still image standards, is the Region of Interest (ROI) coding. Until now, there have been many proposed ROI methods. However, all of these methods can not be applied completely in applications. In this paper, the ROI coding mechanisms in JPEG2000 is first outlined and the diversity of ROI coding methods is discussed in detail. Then a comparative selection about using a proper ROI method according to ROI applications' requirements is proposed. Finally the experimental results show that there are also several parameters that control the effectiveness of ROI coding, such as the codeblock size, ROI size, wavelet filter type, the number of wavelet decomposition levels, and ROI importance score.

1. Introduction

The new image compression standard JPEG2000 has many efficient functionalities. One of its supported is the region-of-interest (ROI) coding scheme [1]. The ROI coding allows different regions of an image to be coded with differing fidelity. To interpret an image is important in applications, where particular regions in an image may be of higher importance than others. Moreover, during the transmission of the image, they need to be transmitted first or at a higher priority. While, because of the multiplex applications of JPEG2000, the requirements of compression ratio have increased. Hence, it is important to have an excellent estimation of proper ROI coding mechanisms to speed up the processing and achieve high requirements with corresponding application.

The ROI coding methods defined in JPEG2000 standard[2][3], as well as several extended ROI coding methods proposed in recent years, are not fully flexible to be useful for diversity of applications. Moreover, a number of ROI coding parameters affect the coding of an image such as the codeblock size, wavelet filter type, number of wavelet decomposition levels etc. Proper selections about ROI coding methods and parameters are very influential to achieve different requirements of various applications.

The remainder of this paper is organized as follows. Section 2 presents an overview of ROI coding mechanisms in JPEG2000. A comparison and a comparative selection of these methods are proposed in section 3, followed by an analysis about ROI coding parameters. Then experimental results will be shown in section 4. Finally, concluding remarks are given in section 5.

2. Overview of JPEG2000 ROI coding mechanisms

ROI coding is capable of delivering high reconstruction quality over user-specified spatial regions in a limited time. In the following, the ROI coding mechanisms will be specified into three main types based on different coding parts of JPEG2000: based on tiling; based on coefficient scaling; based on EBCOT. The coding mechanism based on tiling is not an efficient ROI method, so it will be described in brief. The two other ones, which can be coded statically and dynamically respectively, will be detailedly discussed in the following.

A. Based on tiling

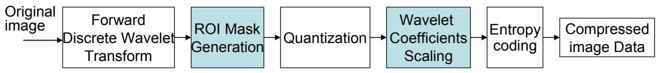
JPEG 2000 allows the spatial partitioning of an image into rectangular and non-overlapping sub-images, called tiles, which can be encoded independently as separate images, to allow access to smaller portions of the image [1] [2]. The tile size can be as small as a single pixel to the size of the original image. The method is particularly desirable in applications where the available memory is limited compared to the image size. However, the rate-distortion performance decreases with decreasing tile sizes, particularly at low bit-rates.

B. Based on coefficients scaling

This ROI mechanism only supports static ROI coding, known as "ROI coding during encoding". It is relevant to applications in which relative importance of different regions in the image is known at the time when the image is compressed. Just as it implies, the main ROI steps are approached during encoding process, as following:

1. Calculating the wavelet transform.
2. If a ROI has been chosen in the image, a transform-domain ROI mask is derived.
3. After quantizing the wavelet coefficients, the quantized coefficients are arranged with a sign-magnitude representation.
4. Using the different of coefficient bit-planes shifting to achieve different ordering techniques and properties.
5. Entropy encoding and transmitting the resulting coefficients progressively with the most important information first. (See Figure 1)

From the most significant bit-plane, plentiful information for the ROI can be encoded foremost, while less for the background areas. The decoding process at the decoder has to reverse the steps above to reconstruct the image. However, with different shifting strategy, the decoding processes may have some different points at decoder.



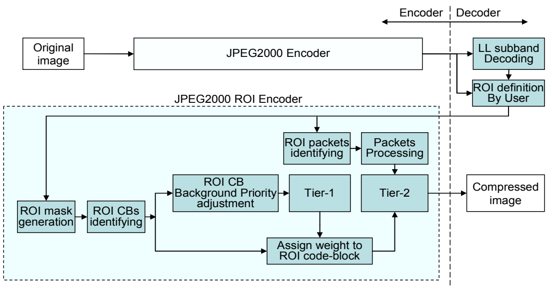
(Fig.1) Static ROI coding mode

The max-shift method (2) and scaled based method (3) are specified in JPEG2000 standard part1 and part2 respectively. They scale the wavelet coefficients contributing to the ROI reconstruction so that the bits associated with ROI are placed in the bit stream before background. With the former, there is must be no overlap between background and ROI bit-planes, and it is efficient for ROI reconstruction. But it has a drawback that it is difficult to control the background important score. The latter, the general scaled method, has a finer control on ROI and background distortion reductions, however, the ROI shape information must be transmitted to the decoder.

Other ROI coding algorithms also exist and have been proposed within JPEG2000 coder in [4-8]. Each algorithm exhibits different features and limitations.

C. Based on EBCOT

This ROI mechanism, supporting both static and dynamic ROI coding, is based on the core coding engine of JPEG2000, the EBCOT algorithm (9). With this mechanism, ROIs can be dynamically defined in interactive environments (Figure 2).



(Fig.2) Dynamic ROI coding mode

Recall that the JPEG2000 code-stream is comprised of a number of quality layers, where each layer comprises of an arbitrary contribution from the embedded bit-stream of each code-block of each sub-band. The implicit ROI coding method which is defined in [9] is used to increase the quality associated with a ROI by including a relatively larger contribution from code-blocks, which are involved in the ROIs reconstruction, in earlier quality layers of the final code-stream. This may be achieved by modulating the distortion estimates of each coding pass, and be input to the PCRD optimization for rate control using the newly defined modulated rate-distortion optimization criterion to drive the formation of the quality layers(9). It is of low complexity in that no bit-plane shifts occur at the encoder or the decoder. However, the ROI adjustments can only be made on a code-block by code-block basis as discussed above.

Several improved ROI coding methods that support dynamic mode are proposed in [10-13], as basic frames showed in Figure 2.

3. Comparisons and a comparative selection of ROI coding methods

In this section, first a comparison of the several ROI methods will be shown detailedly in table1. The details of these algorithms are shown [2-13].

A. Comparisons of ROI mechanisms

As illustrated in table 1, ①② are defined in JPEG2000 standard part1 and part2 respectively as last section showed. ③④⑤⑥ are extended ROI methods, based on coefficients scaling and proposed in recent years. These methods improve two standard ROI methods and integrate their advantages. However, they are not fully compatible with the current JPEG2000 standard except method③. Method③ is suitable for applications where visually lossless reconstruction of ROIs and BG is acceptable because some data about ROI and BG can be deleted in this method.

Method⑦, implicit method, is based on the JPEG2000 coding engine and uses code-block as a coding unit. The modified implicit method⑧ overcomes ⑦'s disadvantages and imports complete/incomplete ROI code-block conception. ① ⑨ ⑪ can not have a flexible control of ROI/BG importance. However, method⑪ can determine how much background information should be included with the ROI packets. Method⑩ is a fast ROI transcoding method for JPEG2000 images. This method takes advantage of the empty packet property to elevate the priority of the ROI with low complexity and faster processing. ⑧⑩ may make some distortion because the truncation of wavelet coefficients and deletion of packets.

<Table1> Summary comparisons of ROI methods in JPEG2000

Coding Part based	ROI coding methods [re.]	Compati-bility	Dynamic/ Static mode	ROI-shape	ROI coding unit	Degrees of ROI importance	Lossy to lossless ROI and BG reconstruction	Control of ROI/BG importance	Published Year
Based on Coefficient Scaling	①Maxshift [2]	Part1	Static	Arbitrary	WC	Same	Yes	No	2000
	②Scaled based [3]	Part2	Static	Rectangle and Ellipse	WC	Different	Yes	Yes	2000
	③Maxshift-like [4]	Part1	Static	Arbitrary	WC	Same	No	Yes	2001
	④BbBShift/GBbBShift [5] [6]	No	Static	Arbitrary	WC	Same	Yes	Yes	2002
	⑤PSBShift [7]	No	Static	Arbitrary	WC	Different	Yes	Yes	2003
	⑥HBShift [8]	No	Static	Arbitrary	WC	Different	Yes	Yes	2004
Based on EBCOT	⑦Implicit ROI [9]	Part1	Dynamic	Regular Polygon	CB	Different	Yes	Yes	2002
	⑧Proposed implicit ROI [10]	Part1	Dynamic	Arbitrary	CB/WC	Different	No	Yes	2005
	⑨Flexible precinct/layer mechanism [11]	Part1	Dynamic	Regular Polygon	Precinct	Same	Yes	No	2002
	⑩Fast ROI Transcoding [12]	Part1	Dynamic	Regular Polygon	Packet	Different	No	Yes	2005
	⑪Prioritized ROI coding [13]	Part1	Dynamic	Regular Polygon	Packet	Different	Yes	No*	2004

B. A comparative selection of ROI coding methods

Generally, based on ROI coding features the ROI coding environments can be classified as non-interactive and interactive environments. Moreover, according to different image applications, the most important ROI requirements are ROI importance score, real-time and multi-ROI. Table2 illustrates a comparative selection about how to choose proper ROI method with different ROI requirements.

<Table2> Comparative selections of ROI methods

ROI requirements	Non-interactive ROI coding	Interactive ROI coding
ROI importance score	①, ②③④⑤⑥	⑨, ⑦⑧⑩⑪
Real-time	③, ① ④⑤⑥, ②	⑨ ⑩, ⑦⑧, ⑪
Multi-ROI	②⑤⑥, ① ③ ④	⑦⑧ ⑩ ⑪, ⑨

ROIs can be coded with the most importance score using ①⑨ since these two methods encode all ROIs at first and cannot control BG/ROI importance. ③⑨⑩ are faster ROI coding methods in a real-time environment despite the truncation of wavelet coefficients and the deletion of packets. While multi-ROI is a significant item since there are often more than one ROI in an image required to be coded, ②⑤⑥ have a better performance in non-interactive environment due to their flexible bit-planes shifting.

C. Effects of ROI coding parameters

A number of ROI coding parameters affect the coding of an image. The following provides a brief overview of the effects of each parameter on ROI prioritized coding.

The support length of the wavelet filter affects the effective size of the ROI. When a larger number of decomposition levels are used, those ROI coefficients in the lower resolution sub-bands can affect an even larger number of samples at full resolution. Regardless of which ROI coding method is used, any coefficient affecting a ROI pixel at full resolution should be included in the importance map specification.

The number of quality layers specifies the number of embedded bit-rates for the progressive encoding of an image. Multiple quality layers should be used if the progressive transmission and reconstruction of ROIs and background is desired. Otherwise a single layer can be used to encode a higher quality ROI at lower bit-rates. Note that the overall coding overhead using multiple layers (even up to 50 layers) is negligible compared to a single layered code-stream. However, this overhead is expected to increase at lower bit-rates [(1)].

JPEG 2000 Part 1 uses code-block sizes that are the same for all sub-bands and resolution levels. It has been shown that the use of smaller code-block sizes decreases the lossless coding efficiency, and so the preferred codeblock size for JPEG 2000 and Max-shift, is 64*64[(9)]. For the implicit method, ROI adjustments can only be made on a code-block by code-block basis. Code-blocks in the lower resolution levels relate to an increasing spatial region. This will require more bits for its prioritization, and thus will result in a slower ROI refinement than that achieved with Max-shift.

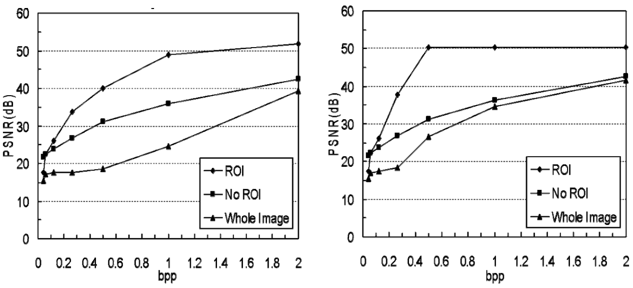
4. Experimental results

Extensive experiments were conducted to test the performance of a number of ROI parameters. Tests were performance on the images in table3. The images used were from the JPEG2000 test set. They are all 8 bits per pixel (bpp) grey-scale, and represent examples from various types of imagery. Image distortion is measured as average peak signal-to-noise ration (PSNR) over all six images. All experiments were completed using version 4.1a of JJ2000, the JPEG2000 (Part5) reference implementation.

<Table3> Experimental images

Image name	Resolution	Type	Image name	Resolution	Type
barb	512*512	Nature	boat	512*512	Nature
mountain	640*480	Nature	satellite	1417 * 793	Aviation
library	464*352	Compound	france	672*496	Compound

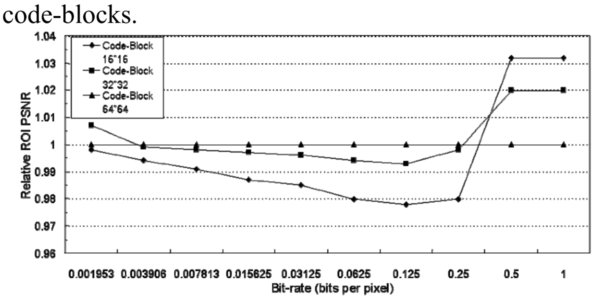
Fig.3 illustrates the effect of reducing the size of the ROI on rate-distortion performance used max-shift method with code-block size 32*32. It can be seen that reducing the size of the ROI decreases the bite rate at which the ROI is received in full detail. The size of the ROI has a complementary effect on the background refinement, as once the ROI has been received in full, code-blocks related to the background will then be present in the bitstream.



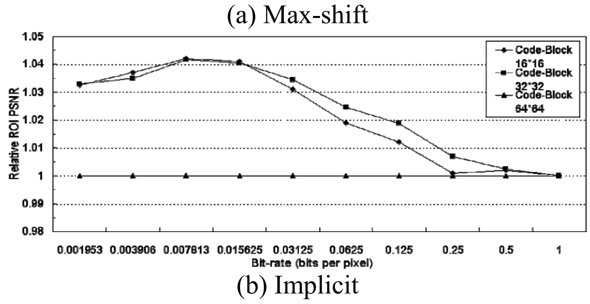
(a) 1/4 ROI size (b) 1/8 ROI size

(Fig.3) ROI PSNR with decreasing ROI size

Fig.4 (a) illustrates the effect of codeblock size on the ROI performance for the max-shift method. The ROI performance using a code-block size of 64×64 is generally superior, especially at low bit-rates. At higher bit-rates (i.e. 0.5 and 1 bpp), the smaller code-block sizes begin to perform better, since their better spatial localization of the ROI more than compensates for the overhead in coding an increased number of code-blocks.



(a) Max-shift

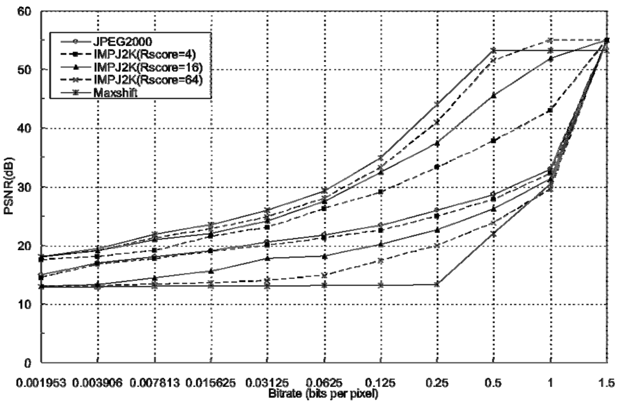


(b) Implicit

(Fig.4) ROI PSNR with decreasing code-block size relative to 64*64

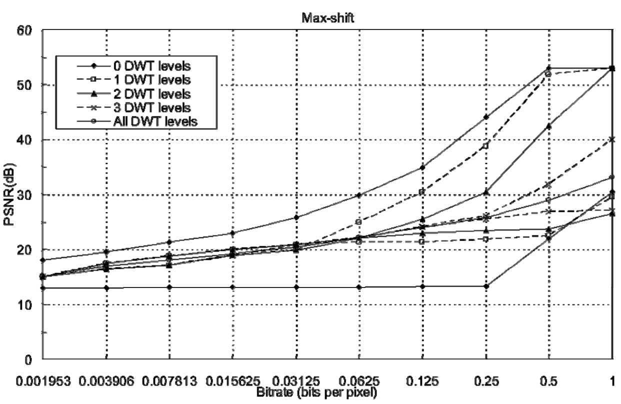
From Fig.4 (b), it can be seen that an increased ROI performance can be achieved with reduced code-block sizes. It can be shown that implicit using 32*32 code-blocks generate code-streams with a larger final bit-rate than that obtained with the max-shift using 64*64 codeblocks. Smaller code-block sizes should be used when the ROI is of primary visual importance at lower bit-rates.

The PSNR performance for a selected number of ROI importance scores are shown in Fig.5. As can be seen, the ROI performance is dependent on the ROI importance score. The larger the importance score, the larger the difference in PSNR between the ROI and background. With JPEG 2000, the ROI and background performances are similar since no ROI emphasis has been introduced into the coding. Max-shift, on the other hand, produces the fastest ROI reconstruction since all ROI bits were encoded before those belonging to the background. The consequence of this is that a very poor background performance results.



(Fig.5) PSNR with different ROI importance scores

Some ROI methods, such as max-shift, implicit methods, provide capabilities for prioritizing some degree of background context. The usual method for achieving this is by applying the same importance score assigned to the ROI to the lower resolution sub-bands of the DWT decomposition. With these ROI coders, increasing the number of low-resolution sub-bands that are to be encoded as part of the ROI will increase the BG quality that will be reconstructed (see Fig.6). The ROI performance, however, will begin to have influence at higher bit-rates. In the extreme case, when all resolution levels are included as part of the ROI, the result is equivalent to a coding scheme with no ROI prioritization.



(Fig.6) PSNR for max-shift with increasing number of lowest-resolution levels

5. Conclusions

In this paper, we have discussed various features and limitations of several ROI coding methods. According to these, a comparative selection of how to choose a proper ROI coding method with different requirements is proposed. It has been demonstrated that to fully utilize the features available in JPEG2000 requires an understanding of both the coding algorithm and the parameters set used to control it.

References

[1] M.Rabbani and R.joshi, "An overview of the JPEG2000 still image compression standard", Signal Processing: Image Communication, vol.17, pp.3-48, 2002

[2] JPEG2000 part1: Final Draft International Standard (ISO/IEC FDIS15444-1), ISO/IECJTC1/SC29/WG1 N1855, August 2000

[3] JPEG2000 part2: Final Committee Draft, ISO/IEC JTC1/SC20 WG1 N2000, December 2000

[4] R.Grosbois, D.Santa-Crus and T.Ebrahimi, "New approach to JPEG2000 compliant Region of Interest coding", Proc. of the SPIE 46th, July 29th August 3rd, 2001

[5] Zhou Wang, Alan C.Bovik and Fellow, "Bitplane-by-Bitplane Shift(BbBShift)-A Suggestion for JPEG2000 Region of Interest Image coding", IEEE signal processing letters, VOL.9,NO.5,May 2002

[6] Zhou Wang, Serene Banerjee, Brian L.Evans and Alan C.Bovik, "Generalized Bitplane-by-Bitplane Shift method for JPEG2000 ROI Coding", IEEE ICIP, 2002

[7] Lijie Liu, Guoliang Fan, "A New JPEG2000 Region-of-Interest Image Coding Method: Partial Significant Bitplanes Shift", IEEE signal processing letters, VOL.10, NO.2, February 2003

[8] Libao Zhang, Ke Wang, "New Approach for JPEG2000 Region of Interest Image Coding: Hybrid Bitplane Shift", Proceedings of the Third International Conference on Machine Learning and Cybernetics, Shanghai, 26-29 August 2004

[9] Taubman, D., and Marcellin, M., "JPEG2000: image compression fundamentals, standards and practice", Kluwer, Norwell, MA, 2002

[10] Hao Yang, Men Long, and Heng-Ming Tai, "Region-of-Interest Image Coding Based on EBCOT", IEE Proceedings-Vision, Image, and Signal Processing, Volume 152, Issue 5, pp. 590-596, October 2005

[11] Rene Rosenbaum, Heidrun Schumanm, "Flexible, Dynamic and Compliant Region of Interest coding in JPEG2000", IEEE ICIP September 22-25, 2002

[12] Kong, H-S;Vetro, A.; Hata, T.; Kuwahara, N., "Fast Region-of-Interest Transcoding for JPEG2000 Images", Mitsubishi Electric Reseerch Laboratories, Inc., 2005

[13] Victor Sanchez, Anup Basu, and Mrinal K. Mandal, "Prioritized region of interest coding in JPEG2000", IEEE Transactions on Circuits and systems for video technology, Vol.14, No.9, September 2004