

## Fast Region-of-Interest Transcoding for JPEG 2000 Images

Kong, H-S; Vetro, A.; Hata, T.; Kuwahara, N.

TR2005-043 May 2005

### Abstract

This paper presents a newly developed region-on-interest (ROI) transcoding method for JPEG 2000 images, which can be used for many progressive image/video streaming applications. The proposed method takes advantage of the empty packet property provided by JPEG 2000 standard to elevate the priority of the ROI. Based on the user specified ROI information and progressive parameters, the developed transcoder extracts the selected ROI and background packets with the desired quality and suppresses the unselected packets to the low quality. Since the transcoder only performs re-organizing the packets in the original coded bitstream, it does not involve any re-encoding operation. Compared to other ROI coding schemes in the literature, the proposed method has the advantages of lower complexity and faster processing. Therefore, it is very suitable for real-time application.

*IEEE International Symposium on Circuits and Systems (ISCAS)*

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.



# FAST REGION-OF-INTEREST TRANSCODING FOR JPEG 2000 IMAGES

Hao-Song Kong<sup>1</sup>, Anthony Vetro<sup>1</sup>, Toshihiko Hata<sup>2</sup>, Naoki Kuwahara<sup>2</sup>

<sup>1</sup>Mitsubishi Electric Research Labs  
201 Broadway, Cambridge, MA 02139, USA  
[hkong@merl.com](mailto:hkong@merl.com)  
[avetro@merl.com](mailto:avetro@merl.com)

<sup>2</sup>Mitsubishi Electric Corporation  
Advanced Technology R&D Center  
Amagasaki, Hyogo 661-8661, Japan  
[Hata.Toshihiko@wrc.melco.co.jp](mailto:Hata.Toshihiko@wrc.melco.co.jp)  
[Kuwahara.Naoki@wrc.melco.co.jp](mailto:Kuwahara.Naoki@wrc.melco.co.jp)

## ABSTRACT

*This paper presents a newly developed region-of-interest (ROI) transcoding method for JPEG 2000 images, which can be used for many progressive image/video streaming applications. The proposed method takes advantage of the empty packet property provided by JPEG 2000 standard to elevate the priority of the ROI. Based on the user specified ROI information and progressive parameters, the developed transcoder extracts the selected ROI and background packets with the desired quality and suppresses the unselected packets to the low quality. Since the transcoder only performs re-organizing the packets in the original coded bitstream, it does not involve any re-encoding operation. Compared to other ROI coding schemes in the literature, the proposed method has the advantages of lower complexity and faster processing. Therefore, it is very suitable for real-time applications.*

## 1. INTRODUCTION

With the growing popularity of mobile wireless communications, there are increasing demands for image and video transmission in various applications. Video surveillance is one of these applications, in which people want to ubiquitously access the stored contents on the server to retrieve important scenes and to remotely browse them quickly. However, due to the network bandwidth limitations, full frame rate and high resolution video transmission are often not allowed. Therefore, a region-of-interest (ROI) transcoding scheme is proposed in the paper to achieve adaptive and scalable video transmission. This scheme is used to extract the important objects (ROIs) from JPEG 2000 coded images, to keep the ROIs with high quality and high resolution and to re-assign low bit rate to the rest of the image or even to truncate the rest of the image based on the bandwidth capacity or user needs.

JPEG 2000 image coding standard [1] provides an excellent hierarchical bitstream structure. It guarantees that the image is compressed once and then the bitstream can be decompressed in many ways to randomly access and extract a sub-set of the bitstream in order to reconstruct the image with the required spatial resolution and quality. In this paper, a detailed method of ROI extraction and bitstream re-organizing is presented. The method is based on the desired features supported in JPEG 2000 standard, such as spatial random access at varying degrees of granularity and degree of quality for each ROI. Compared to other ROI coding methods in the literature, the proposed method has the advantages of lower complexity and faster processing, as it only performs re-organization of the packets in the original coded bitstream without any re-encoding.

## 2. REVIEW OF EXISTING ROI CODING TECHNIQUES

Recently, much attention has been paid to the ROI coding [2]-[6] since the functionality of ROI is suitable for many applications in which certain parts of an image are more meaningful than the other parts of the image. A general method for ROI coding [7] is to scale up the wavelet transformed coefficients of ROIs so that the bits associated with ROIs are more significant than the bits associated with the background (BG). During the image transmission, the ROIs are streamed with higher priorities than the background in a progressive manner. However, depending on the scaling value, some bits of the ROI coefficients may stay on the same bit planes of the background. Therefore, the shape information needs to be transmitted to the decoder to distinguish the ROIs from the background.

JPEG 2000 standard provides a max-shift algorithm [8] to support ROI coding. The max-shift method separates

the ROI from the background by means of scaling them into non-overlapping bit planes. A scaling value is chosen to be sufficiently large to ensure the minimum coefficient associated with the ROI is larger than the maximum coefficient of the background. When the decoder receives the scaling value, it identifies the ROI coefficient by their magnitude. Therefore, the advantage of the method is that it allows arbitrary shaped ROI coding without explicitly transmitting the shape information of the ROI to the decoder. However, this method increases coding overhead due to code blocks on the ROI and background boundary being coded twice.

The major drawback of the scaling based method, including max-shift method, is its static ROI encoding. All ROIs must be defined during the time of encoding. This is a serious problem in the following situations:

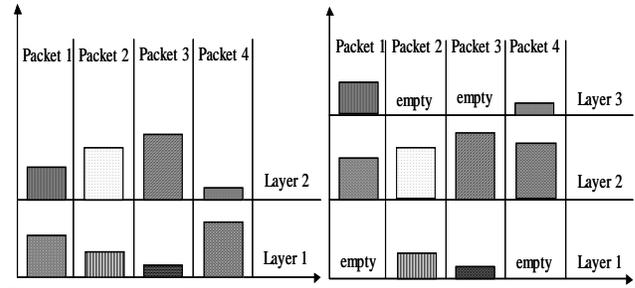
1. the ROI information is not known until the time of decoding
2. the ROI information is provided dynamically by other external sources.

Thus, recently in [9] authors proposed a dynamic ROI coding method, which is able to handle the dynamic ROI information in the interactive environments. The principle of the method is to take advantage of the flexible Packet/Layer mechanism in the JPEG 2000 to rearrange the packet priority in each layer. Since the method follows the layer definition in the original bitstream, it is fully compatible with JPEG 2000 standard. The key technique in the method is its dynamic layer insert functionality. In each layer, the belonging packets are selected from the bitstream according to the ROI information and remain in the same layer; all other packets from non-ROI areas are up-shifted by one layer. In this way, each layer contains only image information from the preferred packets. The resulting non-staffed positions are filled with empty packets. Figure 1 illustrates the procedure of the dynamic layer insertion.

However, the dynamic layer insertion causes the packet header to be re-encoded, since the up-shifted packet is missing the position tag. This, in turn, causes rate-distortion recalculation to generate the information required for the tag-tree encoding for the packet header, which is an undesirable feature for real-time image transmission applications.

### 3. PROPOSED METHOD

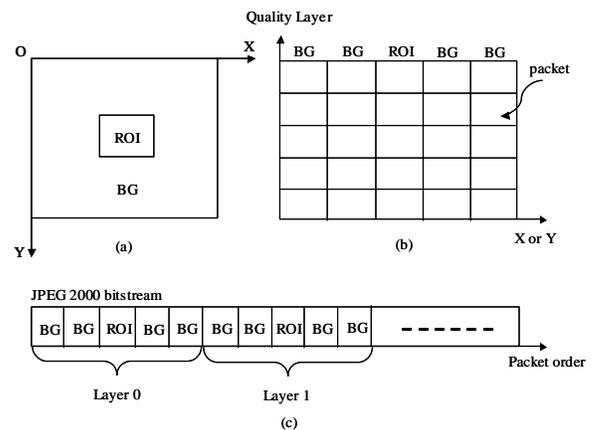
In order to avoid re-encoding of the packet header and make the ROI coding flexible and dynamic, we propose a packet delete (suppress) method in this paper. Instead of pushing the non-ROI packets to the upper layers, the non-ROI packets are deleted from the highest layer down to the



**Figure 1.** (left) The original bitstream contains two layers, each layer has four packets; (right) on layer 1, packets 2 and 3 are selected as ROI, all packets at position 1 and 4 are shifted up and the layer 3 is created for a new layer insertion, the resulting non-staffed positions are filled with empty packets.

lowest layer according to the progressive parameters.

The method makes full use of the empty packet property to effectively elevate the priority of the ROI. This method is simple, easy to implement and performs the entire progressive streaming functionalities based on the ROI information. Figure 2 is used as an example of an image coded with five quality layers to describe our proposed method. It assumes that there exists a selected ROI within the image as shown in Figure 2a, without specifying priorities. The quality layer map for this image is illustrated in one-dimension (first X then Y) as shown in Figure 2b. The corresponding JPEG 2000 bitstream of the image is shown in Figure 2c with Layer-Resolution-Component-Position (LRCP) progression. Due to space constraints, only quality layer structure is shown, the resolution level and component structures are not depicted in the figure. This situation can be thought of as the image containing only one level and one component.



**Figure 2.** An image bitstream representation with LRCP order.

In the proposed method, the transcoding operation is configured with 2 progressive parameters,  $m$  and  $n$ .  $m$  is used to specify the quality of the background and  $n$  is used to specify the quality of ROI. The two parameters must be set with the following restrictions:

If no ROI is selected

$$0 \leq m \leq \text{the maximum quality}$$

else

$$0 \leq m < n \leq \text{the maximum quality}$$

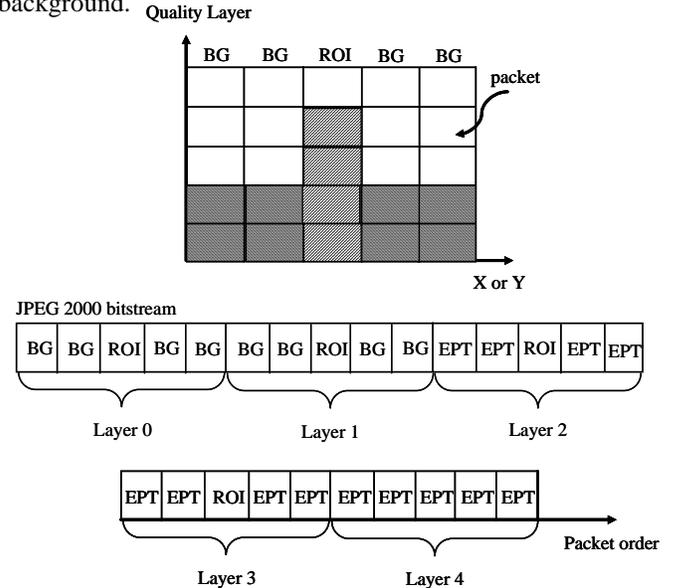
$m=0$  means the packet will be deleted or will be suppressed to the empty packet. The empty packet has a length of zero and only needs one byte for its packet header. The first bit in the packet header should be set to 0 to indicate the empty packet. According to the ROI information and progressive parameters, the transcoder extracts the ROI packet from the lowest quality layer to the quality layer set by  $n$ . The ROI packets on the higher layers are suppressed to the empty packets. For the non-ROI packets, if  $m$  is set to 0, all non-ROI packets are deleted from the highest quality layer to the lowest quality layer. If  $m$  is set not equal to 0, the non-ROI packets are suppressed from the highest quality layer to the lower quality layer set by  $m$ . The non-ROI packets on the layers higher than the layer set by  $m$  are all deleted. An example for  $m=1$  and  $n=3$  for image in Figure 2 is illustrated in Figure 3. The ROI packets are extracted from every layer except for layer 4. The non-ROI packets from layer 0 and 1 are chosen to be rendered but all the packets from the upper layers are suppressed to the empty packets. Other combinations of the progressive parameters will be shown in the figures in the next section.

From the example, it is known that the original bitstream can be tailored or re-organized arbitrarily based on different combinations of progressive parameters to meet the requirements of channel capacity and image quality. The major advantage of the proposed method is that the transcoding operation does not require any re-encoding, even for the re-organized packet headers. All the packets keep their original tag positions during the transcoding, only those unselected packets are suppressed to the empty packets. Therefore, the proposed method performs very fast transcoding. It is very suitable for real-time applications.

#### 4. TRANSCODING EXAMPLES

The proposed method has been applied to many image sequences in a surveillance system. A typical test image "Lena" is used to present the ROI progressive adaptation in this paper. It is assuming that the first requirement for ROI transmission is  $(m,n) = (0,2)$ . Figure 4a shows the extracted ROI with low quality and no background

information. Since most packets are suppressed to the empty packets, the image size is reduced from the original JPEG 2000 coded bitstream 64K to 4K. Figure 4b presents the ROI with  $(m,n) = (0,5)$ . The ROI quality is improved greatly while the background is still suppressed for the sake of saving bits. In the Figure 4c, the ROI with partial background information is shown. Figure 4d and 4e show the high quality ROI with a gradually enhanced background.



**Figure 3.** An example of ROI and background packet extraction and suppression during transcoding.

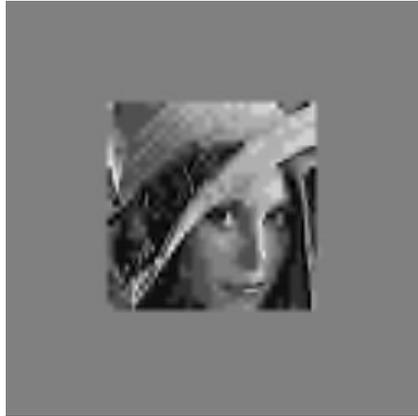
#### 5. CONCLUSIONS

In this paper, the existing ROI coding methods are reviewed. Due to the complexities of the existing methods, a fast and simple transcoding method is proposed. The fast processing is achieved by re-organizing the original bitstream, keeping all the tag positions for every packet and avoiding any re-encoding. The transcoded image examples with various progressive parameters are given to demonstrate the effective progressive transmission.

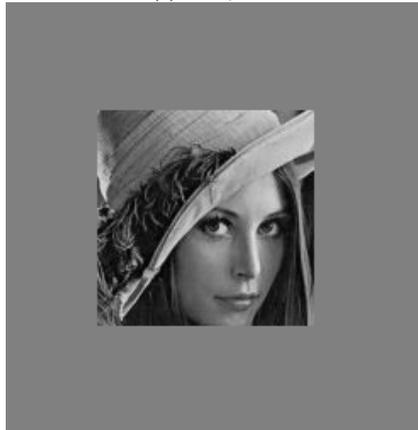
#### REFERENCES

- [1] ISO/IEC 15444-1, "Information technology – JPEG 2000 image coding system – Part 1: Core coding system," 1<sup>st</sup> Ed., 2000.
- [2] R. Grosbois, D. Santa-Cruz, and T. Ebrahimi, "New approach to JPEG 2000 compliant region of interest coding," *SPIE 46<sup>th</sup> Annual Meeting*, San Diego, July, 2001.

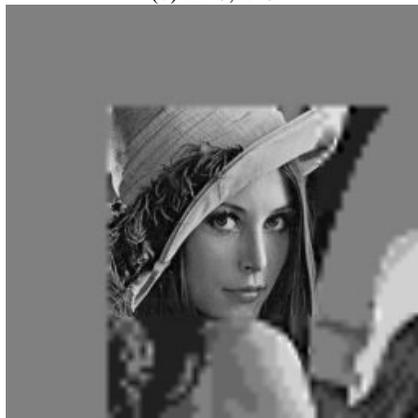
- [3] A. Nguyen, V. Chandran, S. Sridharan, and R. Prandolini, "interpretability performance assessment of JPEG 2000 and Part 1 compliant region of interest coding," *IEEE Trans. Consumer Electronics.*, Vol. 49, No. 4, pp. 808-817, Nov. 2003.



(a)  $m=0, n=2$



(b)  $m=0, n=5$



(c)  $m=0, \text{partial } m=2, n=5$



(d)  $m=1, n=5$



(e)  $m=2, n=5$

**Figure 4.** Progressive transmission.

- [4] L. Liu, and G. Fan, "A new JPEG 2000 region-of-interest image coding method: partial significant bitplanes shift," *IEEE Signal Processing Letters.*, Vol. 10, No. 2, pp. 35-38, Feb. 2003.
- [5] K. Park, and h. Park, "Region-of-interest coding based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 12, No. 2, pp. 106-113, Feb 2002.
- [6] Z. Wang, and A.C. Bovik, "Bitplane-by-bitplane shift (BbBShift) – A suggestion for JPEG 2000 region of interest coding," *IEEE Signal Processing Letters*, Vol. 9, pp. 160-162, May, 2002.
- [7] E. Atsumi and N. Farvardin, "Lossy/lossless region-of-interest image coding based on set partitioning in hierarchical trees," *IEEE ICIP.*, pp. 87-91, Chicago, IL, Oct. 1998.
- [8] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," *IEEE Signal Processing Magazine*, pp. 36-58, Sep, 2001.
- [9] R. Rosenbaum and H. Schumann, "Flexible, dynamic and compliant region of interest coding in JPEG 2000," *IEEE ICIP*, pp. 101-104, Rochester, New York, Sep, 2002.