Transcoding by Automatic ROI Extraction from JPEG2000 Bitstream

Hiroshi Watanabe and Takeo Ogita

Global Information and Telecommunication Institute, Waseda University

1-3-10 Nishi-waseda, Shinjuku-ku, Tokyo 169-0051 JAPAN

email: hiroshi@giti.waseda.ac.jp

Abstract

JPEG2000 supports Region-of-Interest (ROI) for preferential display. In this paper, an approach to extract ROI automatically from JPEG2000 bitstream is considered. By using the ROI, transcoding from the normal JPEG2000 bitstream to the one with ROI is achieved. The validity of the proposed method is confirmed by subjective tests.

1. Introduction

JPEG2000 still image coding [1] provides high coding efficiency as well as lossy to lossless scalability. It is expected to be a major still image coding in the next decade. In the near future, a lot of JPEG2000 bitstreams will appear on many Web sites instead of ones created by the conventional JPEG.

A still image often consists of some meaningful objects and a background. An area in the image where human beings tend to pay attention is called "Region-of-Interest (ROI)." JPEG2000 has a capability to decode and display such ROI area by priority [2]. However, how to determine the ROI area at the encoder is not specified in the standard. A straightforward way to create JPEG2000 bitstream with ROI data is achieved first by extracting ROI from an input image, and then by encoding the image with such ROI information.

Our goal here is to convert normal JPEG2000 bitstream to the one with ROI data. Different from the common approach, we wish to derive ROI form JPEG2000 bitstream or halfway-decoded data for the purpose of transcoding. In this paper, we propose a scheme to transcode JPEG2000 bitsteram without ROI to the one with ROI. It can be achieved by extracting its ROI from wavelet coefficient automatically.

First, we describe a way to extract an arbitrary ROI from an input image. Edge detection, contour tracking and morphological filtering are applied to extract a desired ROI from an input image. An arbitrary shaped ROI can be realized by "MAXSHIFT" method in JPEG2000 without any additional information. Next, the original JPEG2000 bitstream is converted to the one with ROI by this function. A validity of the proposed method is evaluated by comparing the recognition speed of the object in an image in comparison with the normal one. Finally, adaptation to the tiled bitstream is also considered.

2. Basic Concept of JPEG2000 Transcoding

Basic concept of the transcoding is shown in Fig.1. Input JEPG2000 bitsteam is decoded to wavelet coefficients and ROI is extracted from the high pass component. The ROI data is multiplexed to the bitstream.

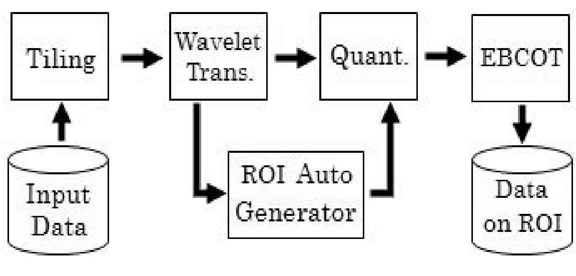
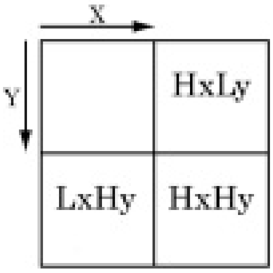


Figure 1. Transcoding from normal JPEG2000 bitstream to the one with ROI.

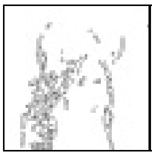
3. Automatic ROI Extraction

To derive ROI from bitstream, it is necessary to carry out halfway decoding to obtain wavelet coefficients. Edge detection is performed to wavelet coefficients, which belong to the highest subbands in horizontal and vertical direction. Simple derivative is enough for this edge detection.

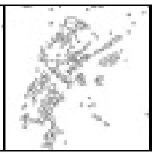
(a) Original



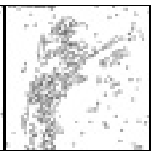
(b) Wavelet Coefficients



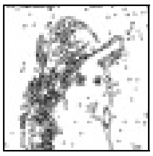
(c) HxLy



(d) LxHy



(e) HxHy



(f) Merged

(g) Background separation

(h) Morphology open

(i) Morphology close

(j) Obtained ROI

Figure 2. Automatic extraction of ROI mask from wavelet coefficient

Obtained edges in some directions are unified and contour is determined based on monitoring histogram of derivatives. Then, morphological filter, which is open and close operation, is applied to cover the ROI and to remove unnecessary isolated objects. As a result, a binary mask can be obtained automatically. This mask represents the area of ROI.

A series of the process is shown in Fig. 2(a)-(j). Fig. 2(a) shows the original image. Fig. 2(b) shows the location of high pass component of wavelet coefficients coded by JPEG2000. Fig. (c)-(e) correspond to contours extracted from modified derivatives of three components of Fig. 2 (b). For this edge detection, absolute difference of the minimum and maximum value within 3x3 pixels area is used for robustness. Next, a histogram of the absolute differences is obtained, and marked pixels are divided into two groups using a threshold since they tend to cluster into high and low values. Marked pixels, which belong to the high value group, are traced to check their connectivity with its neighborhood pixels. Then, outer pixels surrounding obtained traces (Fig. 2(c)-(e)) are merged as Fig. 2(f).

Next, we assumed that ROI locates not around corners of an image but some area close to the center. Thus, we defined pixels from four corners of the image as ones belong to the background. Then, we check pixels value from these four corners in the horizontal and vertical direction. By this operation, an area that is surrounded by marked pixels shown in Fig. 2(f) is also regarded not as the background. Fig. 2(g) shows a pattern traced from four corners of the image of Fig. 2(f). By applying morphology open (Fig. 2(h)) and close operation (Fig. 2(i)) to Fig. 2(g), we can obtain the final mask. These operations are well known to remove isolated small area. Finally, ROI mask shown in Fig. 2(j) can be obtained.

4. Conversion to JPEG2000 Bitstream with ROI

There are two methods specified in JPEG2000 standard to realize ROI function [4]. One is "MAXSHIFT" and the other is "Scale Space." MAXSHIFT can be realized by re-arranging the MSB and LSB of wavelet coefficients in ROI area. Thus, this approach needs to re-calculate quantization and EBCOT, which is entropy coding corresponding to each bit-sliced data. However, "MAXSHIFT" can represent an arbitrary ROI and it does not require any mask information that should be multiplexed in the bitstream. "Scale Space" can only represent a rectangular and an eclipse shape. In addition, data to specify such area should be multiplexed in the bitstream. Therefore, "MAXSHIFT" is adopted in the proposed transcoding. By this approach, the image of the obtained ROI area that depends on the object shape appears first. The rest of them, which is regarded as the background appear later.

5. Experimental Result

We created JPEG2000 bistreams with ROI, which is automatically extracted from the original bitstream. Image resolution is 480x480, coding rate is about 0.5bpp and the file size is 140kByte. For its display, we set the data transmission speed to be 4 KByte/sec between a server and a client [5]. Display process of the image with ROI is shown in Fig. 3 as times go by. Fig. 3 (a), (b) shows 1/10 and 3/4 in a data size. We can recognize an effect of prioritized transmission.

(a) 10% shown

(b) 75% shown

Figure 3 Display process of the image with ROI.

Validity of extracted ROI is evaluated by comparing the recognition speed of the object in an image. We measured the time when viewers recognized the object in an image from the time of starting display. For this experiment, 20 images with/without ROI are used. Each tested image has 500x400 pels in RGB and coding rate is 0.05bpp. Coded file sizes are between 300 and 350 KB. Fig. 4 shows a comparison of time each image is recognized. From this experimental result, it is understood that the proposed approach provides 20% faster recognition time than normal.

6. Adaptation to "Tiling"

If an input image is large, independent coding operation called "tiling" based on separated rectangular block can be used in JPEG2000. The proposed method should be modified to cope with this tiling. A simple way is to extract mask information in each tile and unify the results. Although this approach is quite simple, the obtained results in some experiments show satisfactory results.

(a) Mask

(b) Image

Figure 5 ROI mask for tiled image and result.

7. Conclusion

In this paper, we proposed a new method for transcoding from the normal JPEG2000 bitstream to the one with ROI. The ROI is automatically extracted from the wavelet coefficients of an input image. It is turned out that the proposed

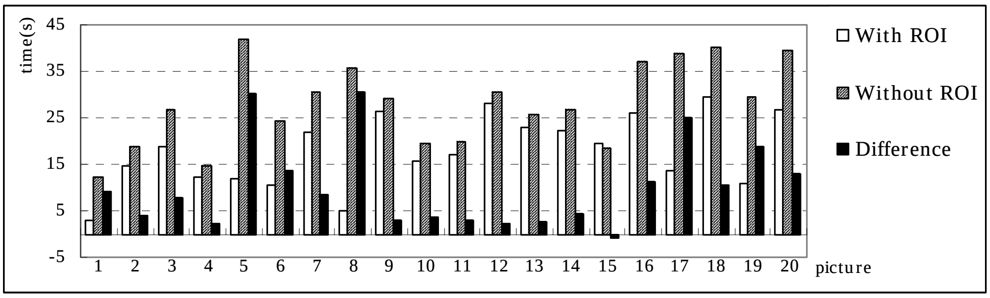


Figure 4. Comparison of recognition time of images.

References

[1] ISO/IEC 15444-1 | ITU-T Rec. 800: "Core coding system," Dec. 2000.

[2] C. Christopoulos, J. Askeloef and M. Larsson: "Efficient methods for encoding regions of interest in the upcoming JPEG2000 still image coding standard," IEEE Signal Processing Letters, Vol.7, No.9, pp.247-249 Sep. 2000

[3] T. Ogita, H. Tatsumi, H. Watanabe: "A study on JPEG2000 transcoder based on automatic ROI extraction," IEICE National Convention, D11-118, Mar. 2002.

[4] Jpeg2000: Image Compression Fundamentals, Standards, and Practice, David Taubman, Michael Marcellin Kluwer Academic Publishers ISBN 079237519X

[5] http://www.kakadusoftware.com