JPEG2000에서 ROI 코딩 파라미터와 ROI 코딩 방법의 상관관계

Correlation of ROI Coding Parameters and ROI Coding Methods in JPEG2000

김호용*, 김형준**, 서영건***

영진전문대학 컴퓨터정보기술계열*, 국방과학연구소 제4체계개발본부**, 경상대학교 컴퓨교육과***

Ho-Yong Kim(hykim@yeungjin.ac.kr)*, Hyung-Jun Kim(khjyis@paran.com)**,
Yeong-Geon Seo(young@gnu.ac.kr)***

요약

JPEG2000은 웨이블릿 기반 정지 영상 압축 표준으로써 앞으로 다양한 분야에서 쓰이게 될 것이다. JPEG2000의 가장 큰 특징 중의 하나는 관심영역(ROI, Region-Of-Interest) 코딩을 제공하는 것이다. 이것은 사용자가 원하는 영역을 다른 영역보다 더 큰 비트율로 압축하여 화질을 좋게 하는 것이다. JPEG2000과 ROI는 다양한 코딩 파라미터를 갖는다. 타일 크기, ROI 크기, 웨이블릿 필터, ROI 모양, 위치, 코드블록 크기, ROI 개수, 이산 웨이블릿 변환 분해 레벨, ROI 중요도, 품질레이어의 수, 저해상도 서브밴드 중요성 등이 있는 데, 본 논문에서는 이들 파라미터와 ROI 코딩 방법 간의 어떤 상관관계를 갖는지 실험을 통하여 보이게 된다. 이는 어떤 응용 프로그램에서 ROI 파라미터와 ROI 코딩 방법을 응용에 맞게 선택하도록 도와준다.

■ 중심어 : | JPEG2000 | ROI | 웨이블릿 |

Abstract

JPEG2000, the standard of still image compression based on wavelet, will be widely used. One of the greatest characteristics of JPEG2000 is to offer ROI(Region-Of-Interest) coding. This is to compress with high quality the region that the user wants better than the other region. JPEG2000 and ROI have different parameters, which are tile size and ROI size, wavelet filter type and ROI shape and its location, codeblock size and number of ROI, number of DWT decomposition level and ROI importance, and number of quality layer and low resolution sub-band importance. In this paper, we shows the correlation of the parameters and ROI coding methods through experiments. This helps an application select the parameters and the methods to meet the application.

■ keyword : | JPEG2000 | ROI | Wavelet |

접수번호 : #060906-001
접수일자 : 2006년 09월 06일
심사완료일 : 2006년 10월 16일

교신저자 : 서영건, e-mail : young@gnu.ac.kr

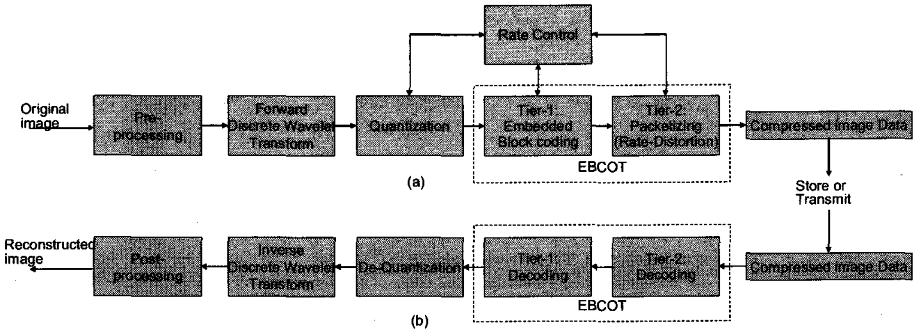

Fig.1. JPEG2000 block diagram of (a)Encoder and (b)Decoder

I. Introduction

One of the most significant characteristics of JPEG2000, the emerging still image standards, is the ROI (Region of Interest) coding. JPEG2000 provides a number of ROI coding mechanisms. Until now on, there are many prioritized ROI coding methods having been proposed, however, all of these methods can not be applied completely and efficiently in applications. The standard, JPEG2000, is intended not only to provide rate-distortion and subjective image quality performance superior to all existing image data compression standards, but also to provide new features and functionalities that current existing standards can either not address efficiently or in some cases can not address at all[1]. One of the efficient functionalities supported by JPEG2000 is the ROI (Region-Of-Interest) coding scheme.

The ROI coding, just as it suggests, allows different regions of an image to be coded with differing fidelity. The functionality of ROI is important in applications where certain parts of the image are of higher important than others. In such a case, these regions need to be encoded at higher quality than the BG (Background, the rest of the image). During the transmission of the image, these regions need to be transmitted first or at a higher priority. The ROI coding methods defined in JPEG2000 standard[2][3], as well as several extended ROI coding methods[4] improved 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 code block size, wavelet filter type, the number of wavelet decomposition levels and etc. Therefore, proper selections about ROI coding methods and the parameters are very influential to achieve different requirements of various applications.

II. Related Works

2.1 JPEG2000 image coding standards

The basic outline of the JPEG2000 encoder incorporates a DWT (Discrete Wavelet Transform) on the source image data, quantization of transform coefficients and then an entropy coding stage before generating the output bit stream[5]. The decoder is the reverse of the encoder, where the encoded bit stream firstly entropy-decoded, de-quantized and then inverse DWT to reconstruct the image data. A more detailed block diagram of the JPEG2000 encoder and decoder are shown in [Fig. 1].

The DWT is first performed on the original input image and generates subbands of wavelet coefficients at a number of resolution levels that describe the horizontal and vertical spatial frequency characteristics of the input image[6][7]. The transform results in four new sub-bands at each level of decomposition, namely, an approximation subband at low resolution, *LL*, and three directionally sensitive detail sub-bands: *LH* - horizontal image features (vertically high pass), *HL* - vertical features (horizontally high pass), and *HH* - diagonal features (horizontally and vertically high pass).

Table 1. Coding parameters of JPEG2000 and ROI

JPEG2000 coding parameters	ROI parameters
- Tile size - Wavelet filter type - Code-block size - # of DWT decomposition level - # of quality layer	- ROI size - ROI shape and location - # of ROIs - ROI importance - Low resolution sub-band importance

The bit-plane coding passes with the highest distortion reduction per average bit of compressed representation should be included in the initial layers. The contributions in any given layer differ from code block to code block, and depend on the distortion (or error) contributions from the coding passes associated with these code blocks. The first quality layer is formed from the optimally truncated code block bit streams such that the target bit rate achieves the highest possible quality in terms of minimizing MSE. Then each subsequent layer is formed by optimally truncating the code block bit streams to achieve higher target bit rates, and thus image quality.

2.2 Coding Parameters

There are several coding parameters both in JPEG2000 and ROI coding having significant effects on ROI coding efficiency. A comprehensive list of the parameters is depicted in [Table 1].

The basic unit to be encoded in JPEG2000 is image tile[1]. An image can be coded as a single tile or can be partitioned into rectangular, non-overlapping, sub-images and each tile coded independently. Tile size is a coding parameter that is explicitly specified in the compressed data. In the JPEG2000 standard, two types of discrete wavelet transform filters are available (see section 1). One is the (9, 7) floating point filter. Another is the reversible (5, 3) integer.

The DWT coefficients are separated into non-overlapping, square regions called code blocks. Code block size is also explicitly specified in the compressed data. JPEG2000 Part 1 uses code block sizes that are the same for all sub-bands and resolution levels. The number of quality layers can

Table 2. The scope of JPEG2000 parameters

JPEG2000 Parameters	Scope (defined in standard)
Tile size	*1~Image size*
DWT filter type	*(5, 3)* and *(9, 7)* filter type
Code block size	$2^n * 2^m$ pixels (n, m>=2, n+m<=12)
Number of DWT decomposition level	*0~32*
Number of quality layer	*1~65535*

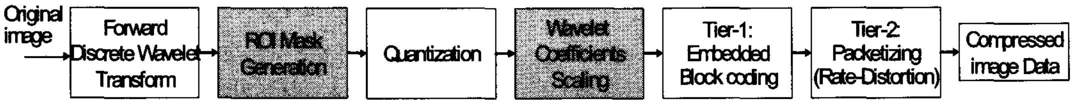

Fig. 2. The encoding processes of static ROI coding

be also specified in JPEG2000. With the increasing of the number of quality layers, the compression performance for SNR progressive applications can be achieved better. Here we summary the scope of parameters defined or recommended by JPEG2000 standard in [Table 2].

III. Factors Influencing the ROI Performance

3.1 ROI Coding Methods

JPEG 2000 provides several ROI coding mechanisms, whereby contents of importance specified by ROIs can be prioritized in the image code stream. According to the different coding algorithms in JPEG2000, we have classified the ROI coding methods into three mechanisms, which are based on tiling, based on coefficients scaling and based on EBCOT[8].

• 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. Tiling produces visible artefacts at the tile boundaries in the reconstructed image, but can be reduced by using a post-processing technique such as adaptive filtering, or the single-sample overlap DWT (SSO-DWT) option offered in JPEG 2000 part 2. Tiling is not an efficient method of ROI coding, unless memory constraints are of primary importance.

• Based on coefficients scaling

This ROI mechanism is the static ROI coding mode, known as "ROI coding during encoding". The concept of coefficients scaling for ROI coding is to shift up/down ROI/BG bitplanes such that coefficients associated with the ROI are placed in higher bitplanes. During the bitplane coding of these coefficients, the ROI will be encoded and placed in the code stream before those associated with the BG. The max-shift[2][9] and scaling based method[3] which are supported in JPEG2000 are based on coefficients scaling. Several other improved methods, such as max-shift like[3][10], PBS[5][12], BbB[11]/GBbB[4], and HBS[6][13] methods, also based on coefficient scaling, have been improved to more efficiently encode the ROI and extend the capabilities of the existing methods.

• Based on EBCOT

This mechanism[14-16] supporting both static and dynamic ROI coding, is based on the core coding engine of JPEG2000, the EBCOT algorithm. With this mechanism, ROIs can be dynamically defined in interactive environments. This mechanism makes use of the EBCOT algorithm (especially tier-2 part) in JPEG2000 by increasing the quality associated with a ROI by including a relatively larger contribution from code blocks or packets, which are involved in the reconstruction of the ROI, into the earlier quality layers of the final code stream.

3.2 Effects of JPEG2000 Coding Parameters

• Tile size

We have known that the tile size can be as small as a single pixel to the size of the original image. However, smaller tiles can reduce the number of decomposition levels in the DWT and this also forces smaller code blocks to be used in the sub-bands that are smaller than the desired code block size. In addition, using tiling at low bit rates can create block artefacts in the images at tile boundaries that significantly detract from the visual quality of the decoded images.

• Code block size

The use of smaller code block size decreases the lossless coding efficiency, and so the preferred code block size for JPEG2000 and Max-shift, is 64×64[4]. The ROI performance using a code block size of 64×64 is generally superior, especially at low bit rates, to the performances of those using reduced code block sizes.

For the implicit method, ROI adjustments can only be made on a code block by code block basis. Since the code block size are to be the same for all sub-bands and resolution levels, code blocks in the lower resolution levels relate to an increasing spatial region. This means that ROI code blocks will contain an increasing spatial extent and will not only relate to the ROI but also regions adjacent to the ROIs.

• Other parameters

The filter type, number of decomposition levels and the number of quality layers can influence the quality. The length of the wavelet filter affects the effective size of the ROI, and this effect is more significant when a larger number of decomposition levels are used.

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

Table 3. Influences on image performance by using coding parameters

Parameters	Block Boundary Artefacts	ROI Performance	ROI Periphery Performance	Background Performance
Tile size	£	£		
DWT filter type		£	£	
Code block size	£	£	£	
Number of DWT decomposition level		£	£	
Number of quality layers		£	£	£
ROI size		£		
ROI shape and location		£		
Number of ROIs		£		
ROI importance		£	£	£
Low resolution sub-band importance score		£		£

ROIs and BG is desired. Otherwise a single layer can be used to encode a higher quality ROI at lower bit rates.

3.3 Effects of ROI Parameters

• ROI size

The JPEG2000 bit streams produced are layer progressive and so increasing rate illustrates the effect of a decoder generating an image of increasing quality as more of the bit stream is received. The ROI size does not adversely affect the code stream bit rate for Implicit. However, for Max-shift, the increase in code stream bit rate is more significant with larger ROIs. This is due to the increase in the number of ROI coefficients with an increased number of bitplanes which have to be encoded.

• ROI shape and location

The effect of ROI shape and location is less important for the Max-shift since ROI encoding is at a coefficient level. Thus, the Max-shift is useful when encoding smaller and/or more complex shaped ROIs. Implicit methods, however, prioritize ROIs at a code block level, and thus have a larger region of influence than the Max-shift. Although an arbitrary shaped ROI may be marked, the reconstruction of ROIs is limited to the spatial region bounded by the code block, sub-band and resolution that contain the ROI. Thus, the performance of the ROI may be improved by choosing a ROI location such that ROI coefficients fall within code block boundaries and only affect a small number of code blocks in the lower resolutions.

• Number of ROIs

Multiple ROIs may be defined for most ROI coding methods, some of which provide the framework and implementations for different importance scores to be assigned to different ROIs. The choice of the number of ROIs affects other ROI parameters such as ROI size and location. Thus, the ROI coding efficiency when using an arbitrary number of ROIs is restricted to that performed by the total combined ROI size and location. For example, if a large number of large ROIs were defined, then an ROI performance can be diminished.

• ROI importance

Some ROI coding methods provide the flexibility to assign arbitrary importance to ROIs to match its 'degree of importance'. This allows a smoother transition from the ROI to BG, as opposed to giving an ROI absolute priority over the BG such as that exhibited by the Max-shift method.

Assuming that the BG has an importance of 1, we can investigate the effect on the performance of the ROI and BG (BG) with increasing ROI importance scores. Importance scores that are powers of 2 were used so that it can be related to the scaling value used in coefficient scaling methods such as Max-shift. The JPEG2000 and Max-shift code streams were also generated to show the two extreme cases of ROI prioritization. JPEG2000 is the case where no ROI prioritization is performed, while the Max-shift using s = 13 (equivalent to an *Rscore* = 8192 (= 2^{13})) refers to the case where all the ROI is prioritized before the BG.

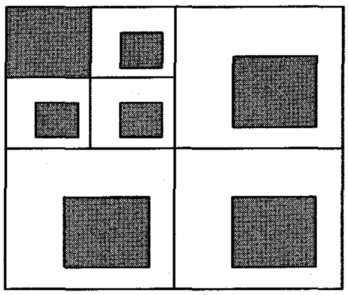

Fig. 3. ROI mask in the lowest resolution level included in the mask

• Low resolution sub-band importance

In some cases, the BG performance can be very poor and may degrade the interpretation of the overall image, especially at low bit rates with high importance. The usual method for achieving this is by applying the same importance assigned to the ROI to the lower resolution sub-bands of the DWT decomposition[9]. An example ROI importance map specification for prioritizing both ROI and low-resolution image information is shown in [Fig. 3].

With the increasing of the number of the low resolution sub-band level taken as ROI, it is possible to get some BG information at an early stage (at low bit rate). [Table 3] shows a conclusion about the influence of coding parameters described in previous chapter on image performance.

IV. Experiments and Evaluations

4.1 Experimental Environments

Extensive experiments are conducted to test the performance of several ROI parameters. Tests are taken on the images in [Table 4]. The images used are from the JPEG2000 test set. They are all 8 bpp grey-scale, and represent examples from various types of imagery.

Table 4. Experimental images

Image name	Resolution	Image name	Resolution
Lena.pgm	512*512	Woman.pgm	2048*2560
Peppers.pgm	512*512	Café.pgm	2048*2560
Barbara.pgm	512*512	Bike.pgm	2048*2560

When the image to be encoded contains an ROI, PSNR is calculated for the ROI alone and over whole image (for the ROI and the BG). All ROI coding is compared to JPEG2000 with its default parameter settings, five level DWT, 64×64 code blocks, 20 layers, layer progressive bit stream etc.

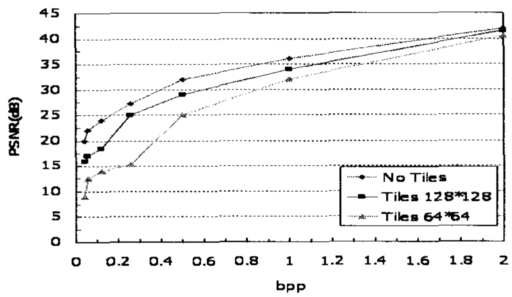

Fig. 4. Rate-Distortion performance with reducing tile size

4.2 Experimental Results of Parameters

[Fig. 4] is clear to observe a decrease in compression efficiency with decreasing tile size. This decrease is particularly significant at bit rate < 1bpp where there is as much as 5dB decrease in PSNR for 128×128 tiles and 10 dB decrease for 64×64 tiles, at the same bit rate, compared to encoding the image as a single tile. It should be noted that the visibility of these block artifacts could be significantly reduced using an adaptive filter as a post-processing operation after decompression. However, this adds significant complexity to the decoder.

[Fig. 5] illustrates the effect of reducing the ROI size on rate-distortion performance. Results presented are for a ROI that is rectangular with a top left hand corner in the centre of the image and size, as a proportion of the total image area, of 1/4 and 1/8. The max-shift algorithm is used to encode the ROI and the code block size is 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, i.e., it increases the speed of ROI refinement.

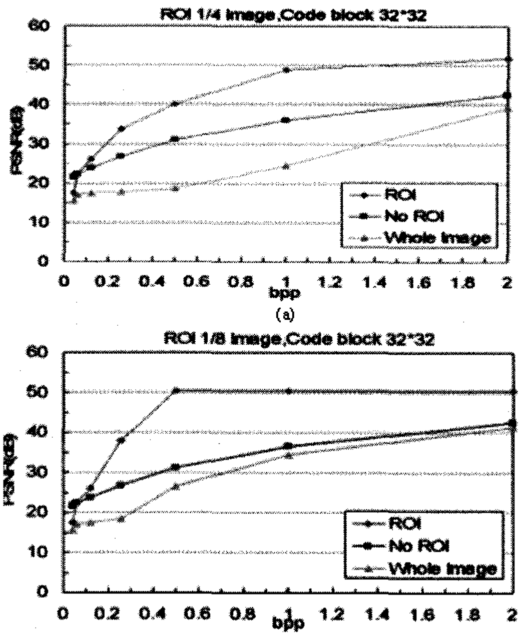

Fig. 5. Rate-distortion performance of ROI 1/4 and 1/8 of image size

When the ROI is 1/4 of the image size, the ROI is not received until the rate is above 1bpp, while the ROI is 1/8 of the image size it has been fully received at 0.5 bpp. This illustrates an approximately linear relationship between ROI size and the rate required to fully decode the ROI. The ROI size has a complementary effect on the BG refinement, as once the ROI has been fully received, code blocks related to the BG will then be present in the bit stream.

[Fig. 6(a)] illustrates the effect of code block 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, to the performances of those using reduced code block sizes. From [Fig. 6(b)], it can be seen that an increased ROI performance can be achieved with reduced code block sizes.

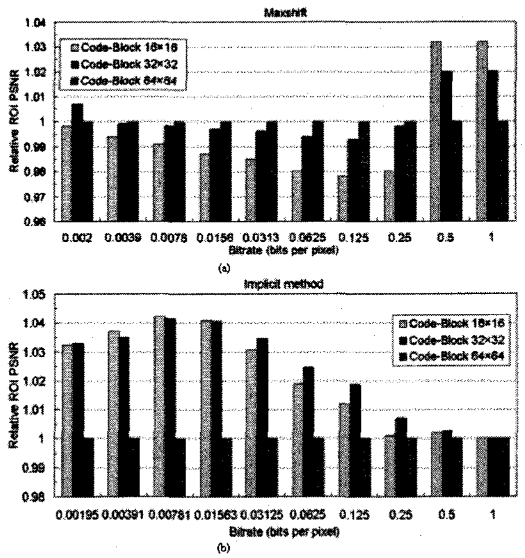

Fig. 6. Relative ROI performance with decreasing code block sizes

The PSNR performance of implicit method for a selected number of ROI importance is shown in [Fig. 7]. The code block size was chosen such that each coding method was not disadvantaged in terms of its rate-distortion performance. A code block size of 64×64 was used for JPEG2000 and Max-shift, while a code block size of 32×32 was used for implicit method.

The implicit method curves are bounded by the two ROI coding extremes, namely, Max-shift and JPEG 2000. As being seen, the ROI performance is dependent on the ROI importance. The larger the importance is, the larger the difference in PSNR between the ROI and BG. That is, an increase in importance increases the PSNR quality of the ROI towards that achieved by the Max-shift, and vice versa for the BG. With JPEG2000, the ROI and BG 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 BG. The consequence of this is that a very poor BG performance results.

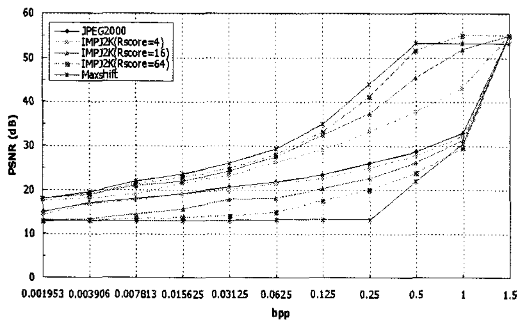

Fig. 7. PSNR performance curves for selected ROI importance

As an experiment. it is clear to decrease in compression efficiency with decreasing tile size. Reducing the ROI size decreases the bite rate at which the ROI is received in full detail, i.e., it increases the speed of ROI refinement In max-shift, the ROI performance using a code block size of 64×64 is generally superior, especially at low bit rates. In implicit, the use of smaller code block sizes allows for increased spatial refinement. The larger the importance is, the larger the difference in PSNR between the ROI and BG.

V. Conclusion

This thesis has presented the research investigation into the field of efficient ROI coding in JPEG2000. JPEG2000 offers significant improvements over previous image compression standards not only in terms of compression performance, but also coding flexibility. However, to fully utilize the ROI feature available in JPEG2000 requires an understanding of both the encoding algorithms and the parameters set used to control it.

References

[1] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 Still Image Coding System: An Overview," IEEE Transactions Consumer Electronics, Vol.46, No.4, pp.1103-1127, Nov. 2000.

[2] ISO/IEC, *International Standard 15444-1, JPEG2000 Image Coding System*, Mar. 2000.

[3] ISO/IEC, *International Standard 15444-2, JPEG2000 Image Coding System: Part II Extensions*, Dec. 2000.

[4] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practic,*. Boston: Kluwer Academic Publishers, 2002.

[5] M. Rabbani and R. Joshi, "An Overview of the JPEG2000 Still Image Compression Standard," Signal Processing: Image Communication, Vol.17, No.1, pp.3-48, Jan. 2002.

[6] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image Coding Using the Wavelet Transform," IEEE Transactions on image Processing: pp.205-220, Apr. 1992.

[7] A. R. Calderbank, I. Daubechies, W. Sweldens, and B. L. Yeo, "Lossless Image Compression Using Integer to Integer Wavelet Transforms," IEEE International Conference of Image Processing, Oct. 1997.

[8] D. S, Taubman, "High Performance Scalable Image Compression with EBCOT," IEEE Transactions on Image Processing, Vol.9, No.7, pp.1158-1170, 2000.

[9] J. Askelof, M. L. Carlander, and C. Christopoulos, "Region of Interest Coding in JPEG2000," Signal Processing: Image Communication, Vol.17, pp.105-111, 2002.

[10] R. Grosbois, D. S. Cruz, and T. Ebrahimi,

"New Approach to JPEG2000 Compliant Region of Interest Coding," Proceedings of the Applications of Digital Image Processing XXIV 4472, pp.267-275, Aug. 2001.

[11] Z. Wang and A. C. Bovik, "Bitplane-by-Bitplane Shift (BbBShift)-A Suggestion for JPEG2000 Region of Interest Image Coding," IEEE Signal Processing Letters, Vol.9, No.5, pp.160-162, May 2002.

[12] L. Liu and G. Fan, "A New JPEG2000 Region-of-Interest Image Coding Method: Partial Significant Bitplanes ShiftM," IEEE Signal Processing Letters, Vol.10, No.2, pp.35-38, Feb. 2003.

[13] L. Zhang and K. Wang, "New Approach for JPEG2000 Region of Interest Image Coding: Hybrid Bitplane Shift," Proceedings of the Third International Conference on Machine Learning and Cybernetics, pp.3955-3960, Aug. 2004.

[14] R. Rosenbaum and H. Schumanm, "Flexible, Dynamic and Compliant Region of Interest coding in JPEG2000," IEEE International Conference on Image Processing, pp.22-25, Sep. 2002.

[15] H. S. Kong, A. Vetro, T. Hata, and N. Kuwahara, "Fast Region-of-Interest Transcoding for JPEG2000 Images," IEEE International Symposium on Circuits and Systems (ISCAS), May 2005.

[16] V. Sanchez, A. Basu, and M. K. Mandal, "Prioritized Region of Interest Coding in JPEG2000", IEEE Transactions on Circuits and Systems for Video Technology, Vol.14, No.9, pp.1149-1155, Sep. 2004.

저 자 소 개

김 호 용(Ho-Yong Kim) 정회원

- 1985년 2월 : 울산대학교 전자계산학과 졸업(공학사)
- 1987년 2월 : 숭실대학교 전자계산학과 졸업(석사)
- 1990년~1991년 : 한국과학기술원 부설 시스템공학연구소
- 1991년~1995년 : 현대자동차(주)
- 1996년~현재 : 영진전문대학 컴퓨터정보기술계열 조교수

<관심분야> : 데이터베이스, 데이터 압축

김 형 준(Hyung-Jun Kim) 정회원

- 1987년 2월 : 부산대학교 계산통계학과 졸업(공학사)
- 1989년 2월 : 숭실대학교 전자계산학과 졸업(석사)
- 1989년~현재 : 국방과학연구소

<관심분야> : JPEG2000, CIM, MPEG21

서 영 건(Yeong-Geon Seo) 정회원

- 1987년 2월 : 경상대학교 전산통계학과 졸업(이학사)
- 1989년 2월 : 숭실대학교 전자계산학과 졸업(석사)
- 1997년 2월 : 숭실대학교 전자계산학과 졸업(박사)
- 1989년~1992년 : 삼보컴퓨터
- 1997년~현재 : 경상대학교 컴퓨터교육과 부교수 경상대학교 부설 컴퓨터정보통신연구소 연구소원

<관심분야> : JPEG2000, Wavelet., ROI