韓國 컴퓨터情報學會 論文誌
第11卷 第4號, 2006. 9.

2006-11-4-2-1

JPEG2000에서 ROI 코딩 품질에 영향을 미치는 요소의 성능 평가

강 기 준*, 이 점 숙*, 김 현 주**

A Performance Evaluation of Factors Influencing the ROI Coding Quality in JPEG2000

Ki Jun Kang *, Jum Sook Lee *, Hyun Joo Kim **

요 약

정지 영상 압축 표준인 JPEG2000의 가장 큰 특징 중의 하나는 관심영역(ROI) 코딩이다. JPEG2000은 다양한 ROI 기법과 ROI 파라미터를 제공하는데, 이것을 특정 응용 프로그램에 적용하기 위해서는 적절한 값을 선택해야 한다. 본 논문에서는 JPEG2000 성능에 영향을 미치는 ROI 코딩 기법과 ROI 파라미터가 ROI 품질과 전체 영상 품질에 어떤 영향을 미치는지를 평가한다. 평가에 사용된 ROI 기법은 Maxshift 방법과 Implicit 방법이며, 평가된 파라미터는 타일 크기와 ROI 크기, 코드블록의 크기, DWT 분해 레벨의 수, ROI 중요도가 사용된다. 타일 크기가 크면 품질은 좋아지며, ROI가 크면 품질은 떨어지며, 코드블록은 Maxshit와 Implicit 모두에서 32X32가 최고의 품질을 보이며, ROI 중요도가 커지면 품질은 떨어지고, 분해 레벨의 수가 증가할수록 품질은 떨어진다.

Abstract

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 and ROI parameters. To apply them to an application, it must select the applicable values. In this paper, we evaluate how the ROI coding mechanisms and the ROI parameters influencing JPEG2000 quality affect the ROI quality and the whole image quality. The ROI coding mechanisms are Maxshift and Implicit, and the parameters are tile size and ROI size, codeblock size, number of DWT decomposition levels and ROI importance. The bigger the tile size, the better the quality. The bigger the ROI size, the ROI importance and the number of DWT decomposition levels, the worse the quality. In code block 32X32 of Maxshift and Implicit, it has the best quality.

▸ Keyword : JPEG2000, ROI, wavelet

• 제1저자 : 강기준 • 교신저자 : 이점숙
• 접수일 : 2006.08.10, 심사일 : 2006.08.28, 심사완료일 : 2006.09.20
* 경상대학교 대학원 컴퓨터과학과 ** 진주산업대학교 컴퓨터공학과 교수
※ 이 논문은 2005년 진주산업대학교 구조개혁 선도대학 지원금 및 간접연구경비를 지원 받아 수행된 연구과제임

Ⅰ. Introduction

With the fast developing multimedia technologies, image compression requires higher performances as well as new features. For example, when processing digital image the main suspense should always be the quality of the image delivered to the receiver. However, the quality should be achieved with as compact representation of the image as possible. To address this requirement in still image encoding, a new standard, the JPEG2000 (Joint Photographic Experts Group 2000), is currently being developed. The standard 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 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[4]. 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.

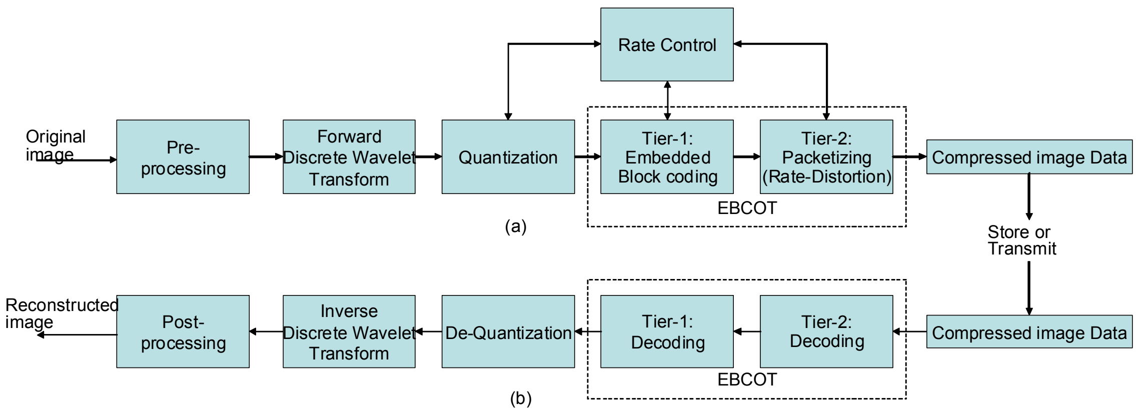

Fig.1. JPEG2000 block diagram of (a)Encoder and (b)Decoder

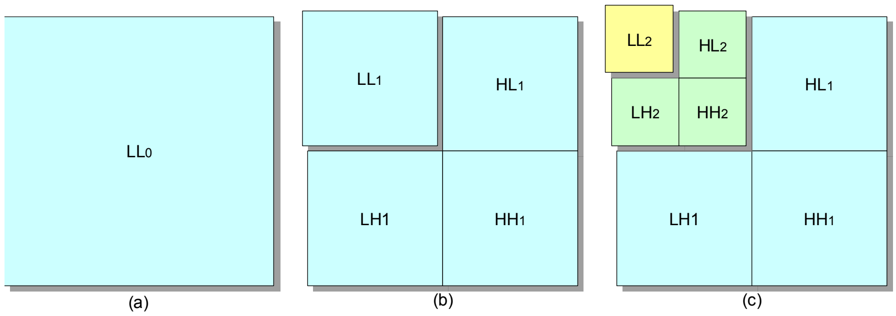

Fig. 2. Example of DWT 2 level decomposition

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[5, 6]. An example DWT is shown in fig. 2. 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).

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 (e.g. layer 1 in fig. 2) 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

Table 1. Coding parameters of JPEG2000 and ROI

JPEG2000 coding parameters	ROI parameters
1) Tile size	1) ROI size
2) Wavelet filter type	2) ROI shape and location
3) Code-block size	3) Number of ROIs
4) Number of DWT decomposition level	4) ROI importance score
5) Number of quality layer	5) Low resolution sub-band importance score

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*

levels. The number of quality layers can 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.

○ 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[7].

○ 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 Maxshift[2,8] and scaling based method[3] which are supported in JPEG2000 are based on coefficients scaling. Several other improved methods, such as Maxshift like[3,9], PBS[4,10], BbB[11]/GBbB[12], and HBS[13,14] 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[15,16,17] 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

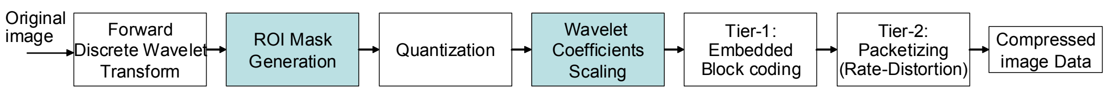

Fig. 3. The encoding processes of static ROI coding

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 Maxshift, is 64×64[5]. 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 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 Maxshift, 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 Maxshift since ROI encoding is at a coefficient level. Thus, the Maxshift 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 Maxshift. 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 Maxshift 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 Maxshift. The JPEG2000 and Maxshift 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 Maxshift using $s = 13$ (equivalent to an *Rscore* = 8192 (= 2^{13})) refers to the case where all the ROI is prioritized before the BG.

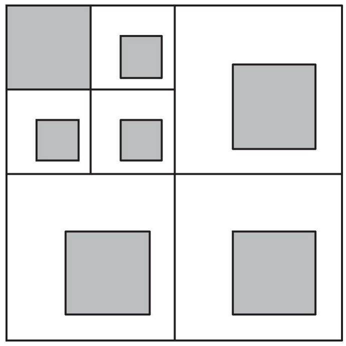

Fig. 4. ROI mask in the lowest resolution level included in the mask

○ 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[8]. An example ROI importance map specification for prioritizing both ROI and low-resolution image information is shown in fig. 4.

3.4 Effects of Several Coding Parameters

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

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		√		√

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.

4.2 Experimental Results of Parameters

Fig. 5 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.

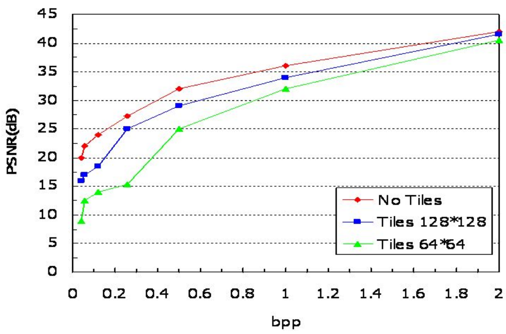

Fig. 5. Rate-Distortion performance with reducing tile size

Fig. 6 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 Maxshift 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.

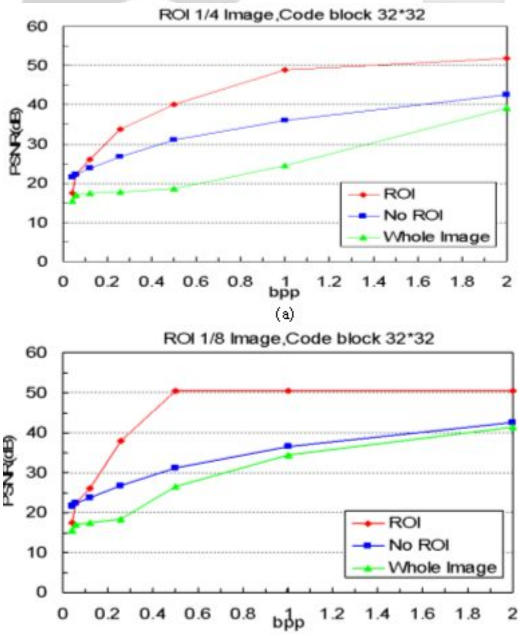

Fig. 6. Rate-distortion performance of ROI 1/4 and 1/8 of image size

When the ROI is 1/4 of the image size, the ROI is not received until the rate is above 1bpp, when the ROI is 1/16 of the image size it has been fully received at 0.25 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 presented in the bit stream.

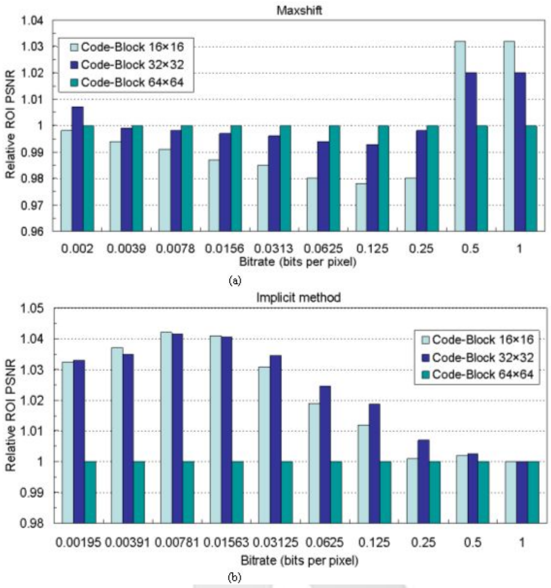

Fig. 7. Relative ROI performance with decreasing code block sizes

Fig. 7(a) illustrates the effect of code block size on the ROI performance for the Maxshift 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. 7(b), it can be seen that an increased ROI performance can be achieved with reduced code block sizes.

The PSNR performance of implicit method for a selected number of ROI importance is shown in fig. 8. 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 Maxshift, while a code block size of 32×32 was used for implicit method.

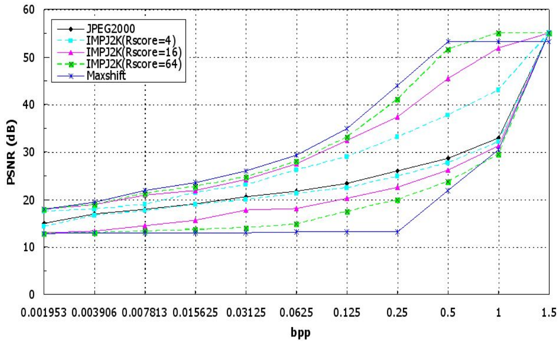

Fig. 8. PSNR performance curves for selected ROI importance

The implicit method curves are bounded by the two ROI coding extremes, namely, Maxshift 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 Maxshift, 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. Maxshift, 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.

With these ROI coders, increasing the number of low-resolution sub-bands that are encoded as part of the ROI raise the BG quality to be reconstructed. In fig. 9, ROI and BG curves are those above and below the ‘All DWT levels’ curve, respectively. Note that ‘0 DWT levels’ represents ROI only coding, and ‘All DWT levels’ represents no ROI coding. 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.

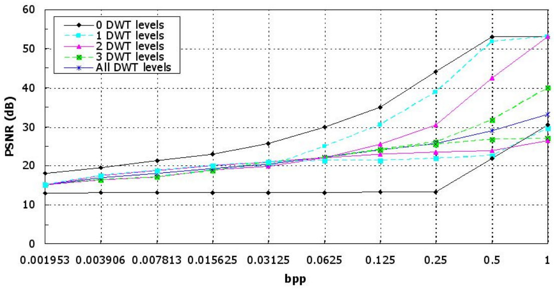

Fig. 9. Rate-distortion with increasing number of lowest resolution levels

참고문헌

[1] Christopoulos C., Skodras A. and Ebrahimi T. Nov. 2000. The JPEG2000 Still Image Coding System: An Overview. IEEE Transactions Consumer Electronics 46(4): 1103 – 1127.

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

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

[4] Rabbani M. and Joshi R. Jan. 2002. An Overview of the JPEG2000 Still Image Compression Standard. Signal Processing: Image Communication 17(1): 3 – 48.

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

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

[7] Bradley A. P. and Stentiford F. W. M. 2002. JPEG2000 and region of interest coding. In Proceedings of the Digital Image Computing Techniques and Applications: 303 – 308.

[8] Askelof J., Carlander M. L. and Christopoulos C. 2002. Region of Interest Coding in JPEG2000. Signal Processing: Image Communication 17: 105 – 111.

[9] Grosbois R., Santa-Cruz D. and Ebrahimi T. Aug.2001. New Approach to JPEG2000 Compliant Region of Interest Coding. In Proceedings of the Applications of Digital Image Processing XXIV 4472: 267 – 275.

[10] Liu L. and Fan G. Feb. 2003. A New JPEG2000 Region-of-Interest Image Coding Method: Partial Significant Bitplanes Shift. IEEE Signal Processing Letters 10(2): 35 – 38.

[11] Wang Z. and Bovik A. C. May 2002. Bitplane-by-Bitplane Shift (BbBShift)-A Suggestion for JPEG2000 Region of Interest Image Coding. IEEE Signal Processing Letters 9(5): 160 – 162.

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

[13] Taubman D. S. 2000. High Performance Scalable Image Compression with EBCOT. IEEE Transactions on Image Processing 9(7): 1158 – 1170.

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

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

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

[17] Sanchez V., Basu A., Mandal M. K. Sep. 2004. Prioritized Region of Interest Coding in JPEG2000. IEEE Transactions on Circuits and Systems for Video Technology 14(9): 1149 - 1155.

저 자 소 개

강 기 준

2000년 경상대학교 컴퓨터과학과 학사
2002년 경상대학교 컴퓨터교육과 석사
2003년~현재 경상대학교 컴퓨터과학과 박사과정
<관심분야> MPEG, JPEG2000, wavelet

이 점 숙

1994년 경상대학교 컴퓨터과학과 학사
1999년 경상대학교 컴퓨터교육과 석사
2003년~현재 경상대학교 컴퓨터과학과 박사과정
<관심분야> MPEG, JPEG2000, wavelet

김 현 주

1988년 경상대학교 전산통계학과 학사
1990년 숭실대학교 전자계산학과 석사
2000년 경상대학교 전자계산학과(공학박사)
1994년~1997년 제일정밀(주) 연구원
2000년~2002년 경남정보대학 컴퓨터정보계열 전임강사
2002년~현재 진주산업대학교 컴퓨터공학과 조교수
<관심분야>: 정보검색, JPEG2000, wavelet