PAPER *Special Section on Papers Selected from the 20th Symposium on Signal Processing*

QoS Estimation Method for JPEG 2000 Coded Image at RTP Layer

Kiyoshi NISHIKAWA[†a)], ***Member***, **Shinichi NAGAWARA**[††b)], ***Student Member***, ***and*** **Hitoshi KIYA**[†c)], ***Member***

SUMMARY In this paper, we propose a novel QoS (Quality of Service) estimation scheme for JPEG 2000 coded image at RTP (realtime transfer protocol) layer without decoding the image. QoS of streaming video is estimated in view of several points, such as, transmission delay, or quality of received images. In this paper, we evaluate the QoS in terms of quality of received images. Generally, RTP is carried on top of UDP, and hence, quality of transmitted images could be degraded due to packet loss. To estimate the quality of received JPEG 2000 coded image without decoding, we use RTP header extension in order to send additional information to the receiver. The effectiveness of the proposed method is confirmed by the computer simulations.

key words: *motion JPEG2000, QoS control, RTP*

1. Introduction

In this paper, we consider the QoS (quality of service) estimation of JPEG 2000 [1] coded image transmission based on RTP (realtime transfer protocol). We propose a method for estimating the QoS (Quality of Service) of the image which is degraded due to the packet loss. The advantage of the proposed method is that it does not require the decoding of the image nor the original image.

Streaming systems, such as video streaming, TV conference, on the Internet demands high quality transmission of motion pictures. For those demanding, one of possible solution is to construct a system based on the Motion JPEG 2000 standard [2], a standard for motion pictures based on JPEG 2000 still image coding [1]. The JPEG 2000 standard provides attractive features, i.e., scalable structure of the code stream, region of interests (ROI), error resilience tools (ERT), and so forth. The proposed method rely on the structure of JPEG 2000 code stream.

RTP [3] is one of the TCP/IP protocol suite for transmitting speech or video streaming. In general, RTP is constructed using UDP (user datagram protocol) as the transport protocol for enabling the low delay transmission. On the other hand, because of using UDP, RTP is not a connection-oriented protocol so that there are no guarantee on the transmission of every packets. In other words, some packet could be lost during the transmission. Therefore, applications using RTP should care the negative effects of packet loss. Hence, it is important to know the quality of transmission, or QoS. The QoS of steaming media, such as motion pictures, is evaluated in several view points, i.e., transmission delay, image quality, and so on [4]. In this paper, we evaluate the QoS of JPEG 2000 image transmission over the Internet by the quality of the received image, namely in terms of peak signal to noise ratio (PSNR) between the original JPEG 2000 encoded image and the received one.

In order to know the exact PSNR, following two steps are required: (1) decoding of the received image, and (2) comparison with the original image. At the receiver side, there are no information on the original image so that the comparison cannot be executed. Moreover, this approach loses the information in the network layer, i.e., the number of lost packets, sequential number of the lost packet and so forth. Without these information, it would be impossible to request the server to retransmit the lost packet. Hence, if there are enough time to wait the packets to be retransmitted we could not send the request to the server.

In this paper, we propose a method to estimate the image quality by sending additional information at the RTP layer. The advantages of the proposed method are that (i) decoding of the received image is not required, and (ii) it provides an estimation of PSNR without the original JPEG 2000 coded image. The proposed method utilizes the structure of JPEG 2000 code stream. In JPEG 2000 coding systems, images are first applied to the discrete wavelet transform (DWT) and then processed after divided in parts by the unit called *code block*. Code blocks are encoded independently from each other so that we can process each code block independently. Thanks to this structure, when bit errors are detected in a code block then the errors can be concealed in the code block. For that purpose, JPEG 2000 standard defines the error resilience tools (ERT). However, even if we use the ERT, PSNR cannot be calculated without the original image.

In the proposed method, we use the RTP header extension for sending additional information in order to estimate the image quality without decoding and the original image. We utilize the information from the RTP header, namely, to locate the lost data due to the packet loss, and to send the information of the original image required to estimate the image quality in terms of PSNR. We are considering to apply the proposed method to the applications, such as image authentication [5], or conference systems in which we re-

Manuscript received December 20, 2005.
Manuscript revised March 15, 2006.
Final manuscript received April 20, 2006.

[†]The authors are with the Course. of Info. and Commn. Sys. Eng., Tokyo Metro. Univ., Hachioji-shi, 192-0397 Japan.
[††]The author is with the Dept. of Electrical Eng., Tokyo Metro. Univ., Hachioji-shi, 192-0397 Japan.
a) E-mail: kiyoshi@eei.metro-u.ac.jp
b) E-mail: nagawara@isys.eei.metro-u.ac.jp
c) E-mail: kiya@eei.metro-u.ac.jp

DOI: 10.1093/ietfec/e89–a.8.2119

ject images which do not reach the predefined desired image quality.

We should note that the proposed method could be regarded as a special case of 'no reference' (NR) scheme to estimate the PSNR of JPEG 2000 images. Some NR methods have been proposed so far [6]–[13] for MPEG, JPEG, or JPEG 2000 coded images. The proposed method differs from those NR methods in the following points. The conventional NR methods estimate the degradation of the image quality due to the image compression such as JPEG, or JPEG 2000. On the other hand, the proposed method estimates the degradation of a JPEG 2000 coded image due to the data loss because of RTP transmission. References [6]–[8] propose to use natural scene statistics (NSS) for quality assessment without the information of the natural (no-compressed) images. However, it is hard to construct a statistical model that describes the effect of packet loss on the image quality and hence, we could not directly apply those NR methods to our problem described above. On the other hand, we send the information extracted from the original JPEG 2000 image using the RTP header extension, and the receiver uses those information to estimate the degradation of image quality due to the packet loss. Hence, we might say that the proposed method is not a pure NR scheme but a special case of it.

This paper is organized as follows. In Sect. 2, we briefly review the structure of JPEG 2000 code stream and the RTP header extension. Then, the proposed method is described in Sect. 3. Section 4 provides results of computer simulation using the proposed method to show its validity.

2. Overview of JPEG 2000 and RTP Header Extension

Here, we will give an overview of the structure of JPEG 2000 images and of RTP header extension.

2.1 JPEG 2000 Coding

In Fig. 1, we show a coding scheme of JPEG 2000. Firstly, an original image is divided into subbands using the discrete wavelet transform (DWT). In the figure, each subband is expressed using a combination of a number and two letters like, '1HL,' or '2LH': The number shows the decomposition level of DWT, and the following letters shows the horizontal and vertical bands, i.e., '1HL' means that the band is obtained by applying high pass filtering horizontally then applying lowpass filtering vertically at the first decomposition level. Wavelet coefficients obtained after the DWT will be divided into code blocks. A code block is the unit for coding the wavelet coefficients.

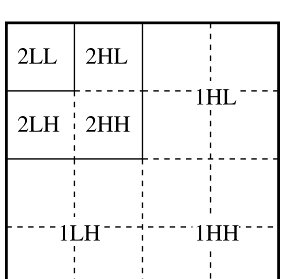

Fig. 1 Division into code blocks based on the wavelet transform.

Each code block will be coded a bit plane at a time from the most significant bit (MSB) to the least significant bit (LSB). Wavelet coefficients in each bit plane are coded in one of the following three coding passes,

- significance propagation (SP) pass
- magnitude refinement (MR) pass
- clean up (CU) pass

and then applied to the arithmetic encoder (Fig. 2).

We should note that each coding pass will be encoded with dependency on its corresponding pass of the upper bit planes. Hence, for decoding a coding pass in a bit plane, data in the upper bit planes are required to decode correctly.

Encoded code blocks are divided into JPEG 2000 packets according to the display level. Each JPEG 2000 packet consists of packet header and body. The packet header contains the information required to decode the data included in its body part which contains a part of the encoded data.

In Fig. 3, we show an example of a JPEG 2000 code stream. The stream starts with the JPEG2000 main header which contains meta information of the image, and it ends with the end of code stream (EOC) mark.

2.2 Error Resilience Tool

In the JPEG 2000 standard, error resilience tools (ERT) are

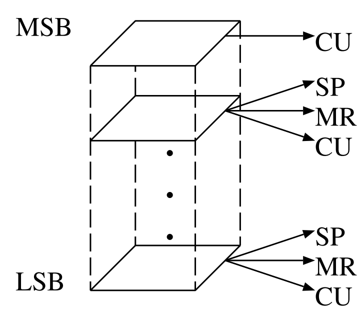

Fig. 2 Three coding passes for each bit plane.

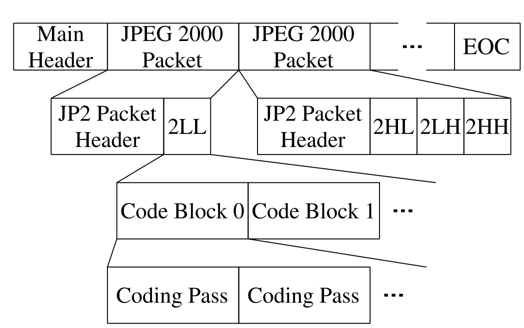

Fig. 3 Structure of code stream of JPEG 2000 coded image.

defined [1], [14] and using the ERT we can detect and conceal bit errors in the code stream. We show some symbols and markers defined in ERT [1] which are related to the proposed method:

- **Segmentation symbols**
  For detecting bit errors in a bit plane at the decoder.
- **Predictable termination** (Predictable)
  For detecting bit errors at the arithmetic decoder.
- **Termination of the arithmetic coder for each pass** (Termination)
  To terminate a coding pass at the arithmetic coder.
- **SOP** (Start Of Packet)
  For finding the start of a JPEG 2000 packet to resynchronize in case of some bits are missing.
- **EPH** (End Of Packet Header)
  For finding the end of the JPEG 2000 packet header to detect errors in the header.

If a bit error occurs in code stream in a pass of a code block, the bits after the erroneous bit can not be decoded correctly, and hence, the quality of the image degrades. By using 'Termination' and 'Predictable' markers of ERT, we can detect the coding pass that cannot be decoded correctly [15]. When we detect a bit error in a code block, then we can delete the coding passes which depend on the erroneous data for avoiding the propagation of errors. We should note that the value of coefficients, which are in the bit plane of the deleted code blocks, will be set to 0 after the processing of the ERT.

2.3 MSE Due to Bit Errors

In this paper, we propose a method to estimate the MSE in the spatial domain using the information on the errors in DWT domain that are caused by the packet loss in the transmission.

Let us assume that bit errors occur in the subband b (one of LL, HL, LH, or HH) at the decomposition level r. The MSE in the spatial domain ε_{sp} due to the errors in the DWT domain can be expressed [16] as

$$\varepsilon_{sp} = \frac{1}{S}\sum_{r=1}^{N_r}\sum_{b=1}^{\nu(r)}\left(Wb_{r,b}^2 \times \sum_{x,y}\left|E_{r,b}(x,y) - W_{r,b}(x,y)\right|^2 \right) \tag{1}$$

where we used the following variables

r	:	decomposition level
b	:	subband (one of LL, LH, HL, or HH)
$W_{r,b}(x,y)$	:	DWT coefficient of original image
$E_{r,b}(x,y)$	:	DWT coefficient of image with errors
S	:	The size of the image [pixel]
N_r	:	The total number of r
$\nu(r)$	:	Number of subband at r

and Wb shows the wighting coefficients determined by the filter coefficients used for DWT. In Table 1, we show the values of Wb when Daubechies 9/7 filter, one of the default filters in the JPEG2000 lossy compression standard, is used. Also, the values of Wb for 5/3 filter are shown in Table 2. We note that Eq. (1) can be used to estimate the MSE regardless of the filters used to implement the DWT, or the number of layers.

Using the values in Tables 1 and 2 with Eq. (1), the MSE can be estimated for the decomposition levels $r = 1$ through 5.

Table 1 Weighting coefficients. (9/7 filter)

Level	LL	HL	LH	HH
5	40.74	20.29	20.29	10.11
4	19.62	9.733	9.733	4.828
3	9.409	4.604	4.604	2.253
2	4.447	2.121	2.121	1.012
1	2.047	1.032	1.032	0.5202

Table 2 Weighting coefficients. (5/3 filter)

Level	LL	HL	LH	HH
5	21.34	11.34	11.34	6.02
4	10.69	5.703	5.703	3.043
3	5.375	2.920	2.920	1.586
2	2.750	1.592	1.592	0.922
1	1.50	1.038	1.038	0.719

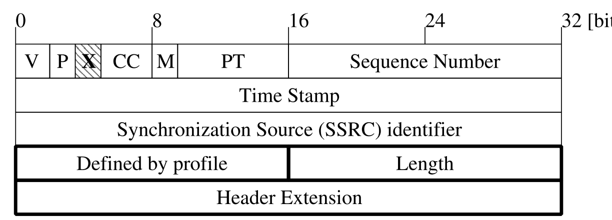

Fig. 4 Structure of RTP header.

2.4 RTP Header Extension

Here, we present a review of the network transmission using RTP and its header extension [3], [17].

When we use RTP, JPEG 2000 code stream of an image is transmitted after they are divided into RTP packets. An RTP packet consists of header and payload parts. The header includes an information on both the network related and the data in the payload, and the payload contains a chunk of JPEG 2000 code stream.

In Fig. 4, we show the structure of the header of an RTP packet. Here, we briefly explain the components of the header that are used in the proposed method. For the explanation of the excluded components, please refer to Refes. [3], [17]. At first, 'Sequence Number' shows the sequential number of the RTP packet for sending. We can use this information for sorting received packets and for detecting the lost packets. 'Time Stamp' shows the time information that is used for synchronization at playback.

The 'X' (eXtensions) flag in the RTP header is for expressing the existence of the header extension. By setting the flag as '1,' we can extend the RTP header for additional information [3]. Namely additional information can be included at the end of the header part of the packet. In Fig. 4, we show the structure of the RTP header when the extension flag is set. The extended contents are added after the normal components of the RTP header, and we marked the extended part by bold lines in the figure.

The extend part starts with 'Defined by profile' (16 bits) which is to identify the extension. Then, there is 'Length Field' (16 bits) to indicate the length of the extended part, namely, we can use $L \times 4$ [bytes] for the extension where L shows the contents of the 'Length Field.' Note that the decoder which is not aware of the extension of the proposed method can skip the extended header by knowing its length from this field.

The proposed method uses the RTP header extension in order to transmit additional information to the receiver for estimating the QoS. The receiver utilize those information to estimate the quality of decoded images.

3. Proposed Method for QoS Estimation

For calculating the MSE using (1), we need the DWT coefficients of the original image $W_{r,b}(x, y)$ and the received image $E_{r,b}(x, y)$. However, $W_{r,b}(x, y)$ is not available at the receiver side, and hence, we need a method to approximate (1) to estimate the image quality. Here, we will describe the proposed method for estimating the QoS in terms of the image quality.

3.1 Outline of the Proposed Method

In the proposed method, we assume the following:

- In the received code stream, there are some data loss due to the packet loss in the network transmission.
- The image is encoded using the ERT with 'Termination' and 'Predictable' symbols. The proposed method estimate the PSNR of the image after the application of the ERT in the decoder.

As noted above, application of the ERT ensures that the DWT coefficients which corresponding to the lost data will be set as zero. Because we can assume that only packet loss can be detected and no bit error could occur at the RTP and the upper level, decoded DWT coefficients will be either zero or the correct value $W_{r,b}(x, y)$ in (1). Based on this property, we propose a method to estimate the image quality without decoding the image.

In Fig. 5, we show the flowchart of the proposed method and the each process can be expressed as the following:

1. Packtization: a JPEG 2000 code stream is divided into RTP packets. (See Sect. 3.2 for detail.)
2. Insert additional information, such as JPEG2000 header, data from code blocks, into the extended part of the RTP header. (See Sect. 3.3)
3. Send RTP packets.
4. Receive RTP packets.

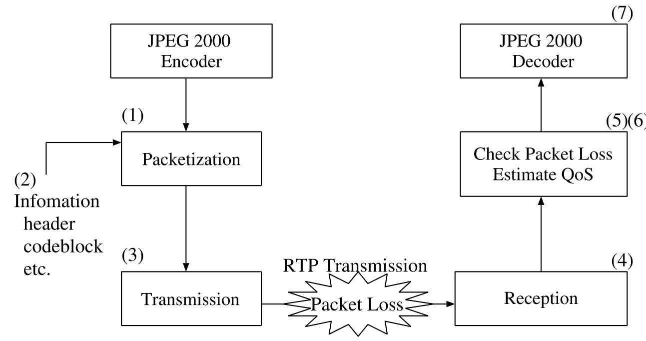

Fig. 5 Flowchart of the proposed method.

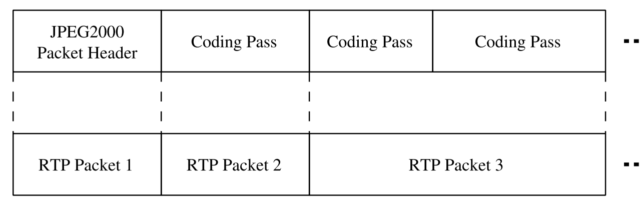

Fig. 6 Decomposition of JPEG 2000 code stream in RTP packets.

5. Check the lost packets using the 'Sequence Number' in the RTP header.
6. Estimating the QoS from the extended header information of RTP packets. (See Sects. 3.4 and 3.6)
7. Pass the reconstructed code stream to the JPEG2000 decoder.

In the proposed method, we use the header extension described in Sect. 2.4 for sending additional information. Those additional information enables the receiver to detect the lost code blocks or lost coding passes. In addition, we can know the number of coding passes that will be rejected by the ERT because of the dependency on the contents of the lost packets. As a result, we can estimate the MSE in the spatial domain using the method described in 2.3.

3.2 RTP Packtization [18]

Here, we explain how to divide code stream into RTP packets in the proposed method. In order to detect the location of lost packets, code stream is segmented according to the structure of JPEG2000 coded image.

In Fig. 6, we show an example of segmentation of JPEG2000 code stream into RTP packets. The rule of segmentation is summarized as below.

1. We prepare dedicated RTP packets for sending JPEG2000 main header, and JPEG2000 packet header respectively.
2. Data of a code block is segmented using a coding pass as a unit. If the size of a coding pass is too small to construct a single RTP packet, then we include several coding passes into one RTP packet.

3.3 Proposed Extension of RTP Header

In the proposed method, in order to locate the data chunk that is contained in a lost packet, we extend the RTP header so that additional information are transmitted. The RTP payload format for the JPEG 2000 image is being considered [19]–[21] and will be standardized in the near future. We should note that the proposed method utilizes only the header extension of RTP, and is independent of the payload format. However, after the standard will be fixed, it could be required to reconsider the information sent to the receiver in the proposed method.

We use three distinct types of additional information, according to the content of the RTP packet as described in Sect. 3.2. They are shown in Figs. 7 through 9 and each component has the following values:

- Common components to the three types
  - M : one bit
    The value '1' indicates that the JPEG 2000 main header is included in the packet.
  - P : one bit
    The value '1' indicates that the JPEG 2000 packet header is included in the packet.
  - S: 1 bit
    The value '1' indicates the packet contains the start of a code block.
  - E: 1 bit
    The value '1' indicates the packet contains the end of a code block.
  - fragment offset:24 bits
    Offset of the data contained in the payload from the beginning of the code stream.
- Items for the packets containing one or more coding passes
  - Code Block: 16 bits
    Indicates the number of code blocks included in the packet.
  - Coding Pass: 8 bits
    Indicates the number of coding passes included in the packet.
  - Total Pass: 8 bits
    Indicates the total number of coding passes in the code block.
- Items for the packet containing the JPEG2000 main header
  - Total A: 32 bits
    A parameter used for estimating the MSE. A detailed explanation would be given in Sect. 3.5.
- Items for the packets containing the JPEG 2000 packet header
  - B: 1 bit
    Set as '1' if the JPEG 2000 packet header contains the information of subbands HL, LH, and HH. Set as '0' if it contains those of subband LL.
  - JP2 Packet No: 8 bits
    Indicates the sequential number of the JPEG 2000 packet header included in the packet.
  - A_{LL}, A_{HL}, A_{LH}, A_{HH}: 16 bits each
    Parameters used for estimating the MSE. A detailed explanation would be given in Sect. 3.5. When 'B' is set as '0' then only 'A_{LL}' is used. When 'B' is set as '1' then 'A_{HL},' 'A_{LH},' 'A_{HH}' are used.

0	8	16	24 – 32 [bit]
M P S E	fragment offset		
CodeBlock		Coding Pass	Total Pass

Fig. 7 Structure of the extended part of RTP header for the packet containing data of coding passes.

0	8	16	24 – 32 [bit]
M P S E	fragment offset		
Total A			

Fig. 8 Structure of the extended part of RTP header for the packet containing JPEG 2000 main header.

0	8	16	24 – 32 [bit]
M P S E B	fragment offset		
JP2 Packet No		A_{HL} (A_{LL})	
A_{LH}		A_{HH}	

Fig. 9 Structure of the extended part of RTP header for the packet containing JPEG 2000 packet headers.

Using these information, we can know the lost portion of code stream if packet losses occur. Namely, if nth packet is lost, we can get the detailed information from the RTP packet headers of $n-1$st and $n+1$st packets. Those obtained information includes the location of the lost code blocks or number of coding passes that can be decoded.

3.4 Proposed Estimation Method of MSE

In the proposed method, we calculate the errors in DWT domain using the additional information in the extended RTP header.

The MSE ε_{sp} of the received image is estimated based on (1) using the information sent by extended RTP headers. In Eq. (1), we should recall that $E_{r,b}(x,y)$ is either 0 or $W_{r,b}(x,y)$ when we consider the transmission in the Internet. With considering this fact, we propose the following formula for obtaining the estimated MSE $\tilde{\varepsilon}_{sp}$:

$$\tilde{\varepsilon}_{\text{sp}} = \frac{1}{S}\left\{\sum_{r=1}^{N_r}\sum_{b=1}^{\nu(r)} A(r,b) - \sum_{r=1}^{N_r}\sum_{b=1}^{\nu(r)} A(r,b)\frac{1}{N_{cb}(r)}\sum_{i_{cb}=1}^{N_{cb}(r)}\sigma(i_{cb})\right\} \quad (2)$$

where N_r, $\nu(r)$ and i_{cb} show the decomposition level, number of code blocks in a subband, and i-th code block respectively. Further, $A(r,b)$ and $\sigma(i_{cb})$ is given as

$$A(r,b) = Wb_{r,b}^2 \times \sum_{x,y} W_{r,b}(x,y)^2 = Wb_{r,b}^2 \times \gamma_{r,b} \quad (3)$$

$$\sigma(i_{cb}) = \begin{cases} \frac{1-0.5^{N_{dec}(i_{cb})}}{1-0.5^{N_{cp}(i_{cb})}} & N_{dec}(i_{cb}) \neq 0 \\ 0 & N_{dec}(i_{cb}) = 0 \end{cases} \quad (4)$$

where $\gamma_{r,b}$ is given as

$$\gamma_{r,b} = \sum_{x,y} W_{r,b}(x,y)^2 \quad (5)$$

and the values of $\gamma_{r,b}$ depends on the DWT coefficients of the processed image. Note that $\sigma(i_{cb})$ shows the contribution of the code block to the quality of the received image. Similarly, $A(r,b)$ shows the contribution of the subband to the image quality.

The variables used in (4) are as follows

$\gamma_{r,b}$	:	the sum of squared DWT coefficients of b-th subband of the r-th decomposition level
$N_{dec}(i_{cb})$	:	number of coding passes in code block i_{cb} that can be correctly decoded
$N_{cp}(i_{cb})$	:	number of coding passes in code block i_{cb}.

In order to estimate the MSE of a decoded image using (2), we need the values of S, N_r, $N_b(r)$, $N_{cb}(r)$, $N_{dec}(i_{cb})$, $A(r,b)$, and $N_{cp}(i_{cb})$. Of these, we can obtain S, N_r, $N_b(r)$, and $N_{cb}(r)$ form the JPEG2000 main header.

On the other hand, $N_{cp}(i_{cb})$ is known from 'Total Pass' in the extended header information (See Fig. 7). We can, then, calculate $N_{dec}(i_{cb})$ from 'coding pass' , 'S,' and 'E' in the extended RTP header. If there are no lost portion in the data of the i_{cb}-th code block, we set $N_{dec}(i_{cb}) = N_{cp}(i_{cb})$.

3.5 Estimation of $A(r,b)$

For calculating $\tilde{\varepsilon}_{\text{sp}}$ using Eq. (2), the values of $A(r,b)$ are required, and they are calculated using DWT coefficients. In the proposed method, we send those information using the extended RTP header. Namely, we use 'Total A,' 'A_{LL},' 'A_{HL},' ''A_{LH}' and 'A_{HH}' in the header for this purpose.

The sender set the values of 'Total A,' 'A_{LL},' 'A_{HL},' 'A_{LH},' and 'A_{HH}' according to the following:

- Total A
  This value shows the total sum of $A(r,b)$, that is,

$$A_T = \sum_{r}^{N_r}\sum_{b}^{\nu(r)} A(r,b)$$

- A_{LL}
  This value shows the sum of squared values of the DWT coefficients, which are contained in the JPEG 2000 packet, multiplied by $Wb_{r,LL}$, that is,

$$A_{LL} = \gamma_{r,LL} \times Wb_{r,LL}^2$$

  where r shows the decomposition level of the JPEG 2000 packet. This value is calculated when 'B' is set as 0.
- A_{HL}
  Calculated using the coefficients of HL subband as

$$A_{HL} = \gamma_{r,HL} \times Wb_{r,HL}^2$$

  for r-th decomposition level of the JPEG 2000 packet. This value is calculated when 'B' is set as 1.
- A_{LH}
  Calculated using the coefficients of LH subband as

$$A_{LH} = \gamma_{r,LH} \times Wb_{r,LH}^2$$

  for r-th decomposition level of the JPEG 2000 packet. This value is calculated when 'B' is set as 1.
- A_{HH}
  Calculated using the coefficients of HH subband as

$$A_{HH} = \gamma_{r,HH} \times Wb_{r,HH}^2$$

  for r-th decomposition level of the JPEG 2000 packet. This value is calculated when 'B' is set as 1.

Note that the values of A_b become real values. For including those values into the RTP header, we should express them in binary format. For that purpose, we convert them into integer using

$$\bar{A}_b = \text{round}\left\{(2^{16}-1) \times A_b\right\} \quad (6)$$

where round(·) shows the rounding operation, and then, $\bar{A}_b$ are transmitted to the receiver. The receiver converts $\bar{A}_b$ into A_b using the inverse operation of Eq. (6). To preserve the accuracy of A_b we use 16 bits for storing each $\bar{A}_b$.

In the proposed method, the sender should obtain these values before the code stream will be sent using RTP. In order to obtain the DWT coefficients, it is required to decode the image, or we should make the JPEG2000 encoder to pass those information.

3.6 Proposed Algorithm for Calculating $\tilde{\varepsilon}_{\text{sp}}$

After receiving the RTP packets, the following process will carried in order to estimate the QoS of the image:

1. Check if the JPEG2000 main header is received.
   If the main header was lost during the transmission, we quit the processing on that image because we cannot decode without the main header information.
2. If the main header is received, the following processes

are applied to each JPEG 2000 packet.

a. Check if the JPEG2000 packet header is received. If the JPEG2000 packet header is lost due to the network packet loss then skip to the processing of the next JPEG2000 packet.
b. Calculate the contribution of the code block for quality of image.
Using (4), we can calculate the contribution $\sigma(i_{cb})$ of the code block using the number of coding passes that can be decoded correctly.
c. By averaging $\sigma(i_{cb})$ of all the code blocks in the JPEG 2000 packet, we get

$$C_b(r,b) = \frac{1}{N_{cb}(r)} \sum_{i_{cb}=1}^{N_{cb}(r)} \sigma(i_{cb}) \tag{7}$$

and which can be considered as the contribution to the image quality of a received DWT coefficients of a subband.
d. Calculate $A(r,b) \times C_b(r,b)$ and summed up for all the subband ('A_{LL},' 'A_{HL},' 'A_{LH},' 'A_{HH}') gives

$$A(r) = \sum_{b=1}^{\nu(r)} A(r,b) C_b(r,b) \tag{8}$$

and this $A(r)$ can be considered as the contribution of a JPEG 2000 packet to the image quality.

3. Finally, we can calculate the estimated MSE $\tilde{\varepsilon}_{\text{sp}}$ according to

$$\tilde{\varepsilon}_{\text{sp}} = \frac{1}{S}\left\{A_T - \sum_{r=1}^{N_r} A(r)\right\}. \tag{9}$$

3.7 Consideration of Network Overhead of the Proposed Method

Here, we consider the overhead on the network traffic caused by the proposed method.

Using the proposed method, the length of each RTP header increases by eight to twelve bytes (See Figs. 7 through 9). This seems to generate the network overhead. Let us consider the overhead, which the proposed method generates, in terms of the number of network packet to transform a JPEG 2000 image.

The Internet is a packet switching network, and hence, we could say that network load is not affected if the number of transmitted network packet does not increase. Each network has its maximum size of data to transmit as one network packet, and this unit for transmission is called as MTU (maximum transmission unit). When the size of a network packet exceeds the minimum MTU of the network then the packet will be fragmented into several packets of smaller size.

If the size of an RTP packet exceeds the MTU due to the proposed method, then the packet will be fragmented, and hence, the number of transmitted packets increases. In this case, we should say that the proposed method introduces the network overhead.

However, in general, when UDP is used as the transport layer protocol, the size of a packet is selected to be much smaller than the minimum MTU on the path to the destination to prevent the occurrence of packet fragmentation. In this case, we could say that it is rarely happen that the size of a packet exceeds the MTU even if the length of the RTP header increases by eight to twelve bytes. Therefore, we consider that the network overhead due to the proposed method can be neglected when the size of each packet is much smaller than the MTU as is the general case.

4. Simulation Using the Proposed Method

Here, we show results of simulation using the proposed method to demonstrate its effectiveness.

4.1 Conditions for Simulations

We implemented the proposed method for transmitting still images. Three different images, namely, Lena, Barbar, Baboon, were transmitted using the proposed method.

The conditions of the simulation are as following:

- Still images: Lena, Barbara, Baboon
  - Image size: 512×512 [pixel]
  - Number of component: 1
  - Number of Wavelet decomposition: 5
  - Number of layer: 1
  - Code block size: 64×64
  - Used ERT markers: 'Termination of the arithmetic coder for each pass,' 'Predictable termination,' 'SOP'
- RTP payload size: 1024 [bytes]
- Packet loss rate: 5%, 10%, 15%
- JPEG 2000 decoder: Kakadu

We used identical coding conditions during the simulations. Independent thirty trials were done for each image. We calculated peak signal to noise ratio (PSNR)

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \tag{10}$$

from the estimated MSE using the proposed method.

We also calculated the PSNR between the original JPEG 2000 coded image and the decoded one using ERT of JPEG 2000 standard. This calculated PSNR was compared with the estimated PSNR using (10) of the proposed method.

4.2 Results of Simulation

We show the results of simulation using Lena, Barbara, Baboon in Figs. 10, 11, and 12 respectively. In these figures, the horizontal axis shows the number of trial, and the vertical axis shows the PSNR [dB]. Note that results shown in

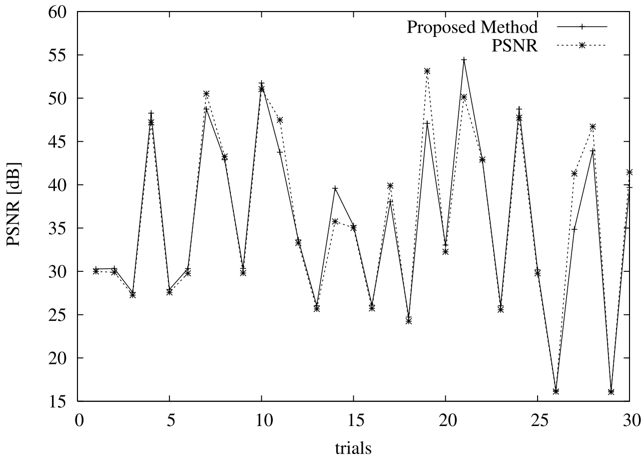

Fig. 10 Results of simulation. (Lena)

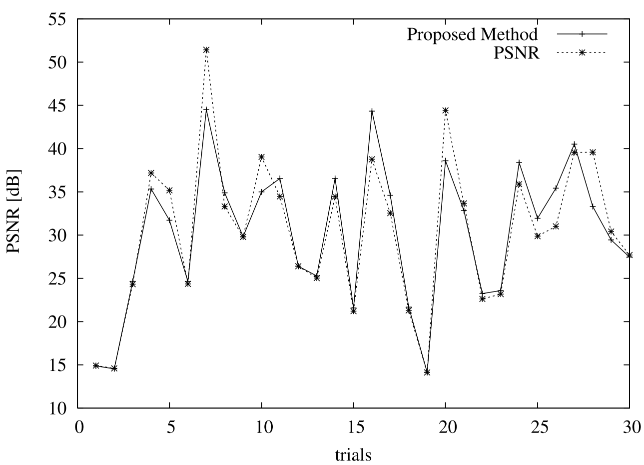

Fig. 11 Results of simulation. (Barbara)

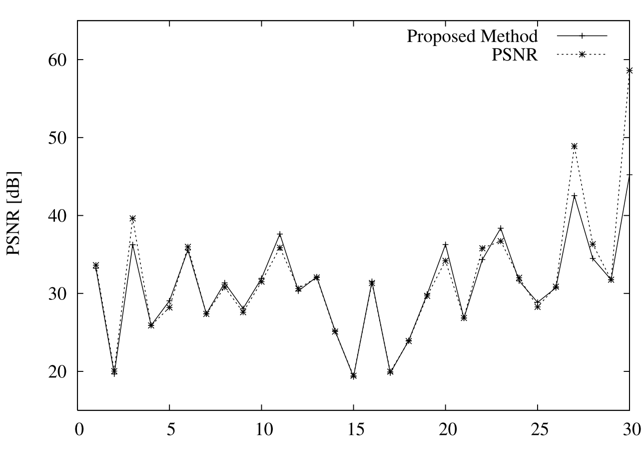

Fig. 12 Results of simulation. (Baboon)

Figs. 10 through 12 are under the condition that the packet error rate was 5%.

From Fig. 10, in case of image Lena, we can see that there is maximum 6 [dB] difference in PSNRs in 11th, 14th, 21st, and 27th trials. The reason of this difference is due to the formula for the estimation in the proposed method. As can be seen from Eq. (2) we express the contribution of each subband as Eq. (3). Also we consider the contribution decreases when some code blocks of a subband are lost due to the packet loss, and the ratio of the decrement is assumed to be in proportional to the ratio of the number of correctly decoded code blocks to the total number of code blocks. This assumption gives a good estimation when its contribution to the image quality is constant over the subband or when the subband is composed of one code block. However, when a subband is divided into several code blocks, and its contribution to the image quality varies depending on the code blocks, we recognize the accuracy of the estimation of the proposed method is slightly decreased. We can say that the same tendency can be seen from Figs. 11 and 12.

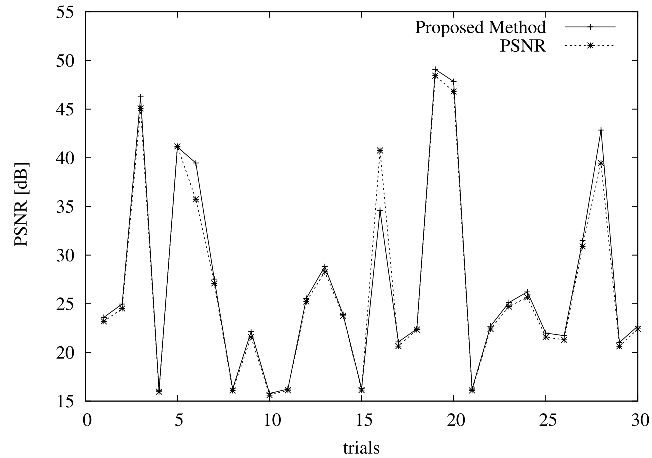

Fig. 13 Results of simulation (Lena) when the packet error rate was 10%.

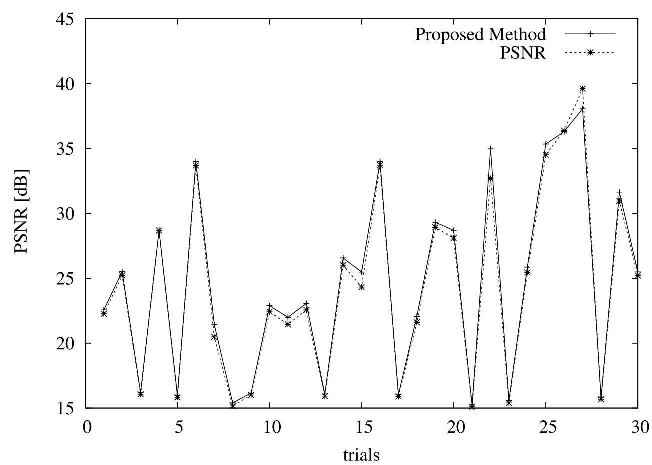

Fig. 14 Results of simulation (Lena) when the packet error rate was 15%.

In order to show the robustness of the proposed method to the packet error rate, we show the results of simulations when error rate was set as 10% and 15% in Figs. 13 and 14. Note that the transmitted image was Lena for both cases. From these figures, we can confirm that the proposed method provides a good estimation of PSNR regardless of the packet error rate.

4.3 Consideration on the Effect of the Approximation of DWT Coefficients

Here, we provide the simulation results for showing the effect of the approximation of the DWT coefficients. Namely, in this case, the values of 'Total A,' 'A_{LL},' 'A_{HL},' 'A_{LH},' 'A_{HH}' in the RTP header are those of image 'Lena' for processing the images 'Barbara' and 'Baboon.' The results are shown in Figs. 15 and 16 for 'Barbara' and 'Baboon' respectively.

By comparing Figs. 11 and 15, and Figs. 12 and 16, we could confirm that the estimation errors are increased. Reason of this degradation is using the DWT coefficients of other image. However, we can see that we can obtain the rough estimation of the quality of image even if we use the DWT coefficients of other images. This result suggest that we can decrease the amount of information to send to estimate the image quality. That is, if motion pictures are sent using the proposed method, frames in succession might similar and hence, we can use the same values for A_T and so on to multiple frames with slight degradation of estimation.

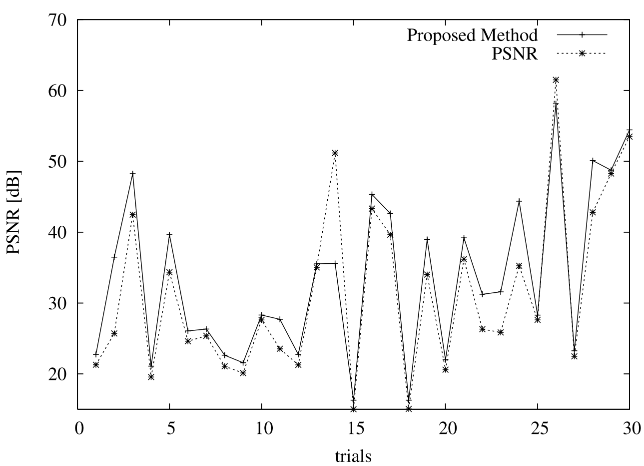

Fig. 15 Results of simulation. (Barbara)

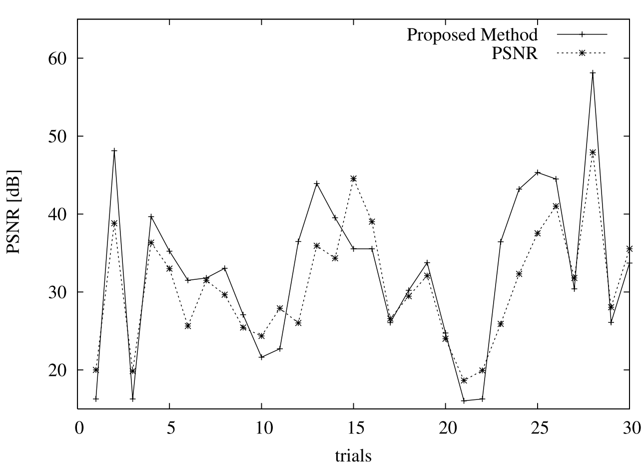

Fig. 16 Results of simulation. (Baboon)

5. Conclusions

In this paper, we proposed a method for estimating the QoS of JPEG 2000 coded image before the processing of the decoder. The effectiveness of the proposed method was confirmed by the computer simulations. As a future work, we will consider applying the proposed method to the multi-layered images.

References

[1] I.T.R.T. ISO/IEC 15444-1, "Information technology—JPEG 2000 image coding system—Part1: Core coding system," ISO/IEC JTC 1/SC 29 WG1, Jan. 2001.

[2] I. 15444-3, "Information technology—JPEG 2000 image coding system—Part3: Motion JPEG 2000," ISO/IEC JTC 1/SC 29 WG1, July 2002.

[3] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A transport protocol for real-time applications," ISO/IEC JTC 1/SC 29 WG1, July 2002.

[4] S. Tasaka, "User-level qos assessment of multimedia transmission over ip networks by qos mapping," Proc. IEICE Gen. Conf. 2005, DS-1-3, March 2005.

[5] H. Kobayashi, S. Nagawara, K. Nishikawa, and H. Kiya, "Image authentication of JPEG 2000 coded images using extended RTP protocol," Proc. IEICE Gen. Conf. 2005, DS-1-4, March 2005.

[6] H.R. Sheikh, A.C. Bovik, and L.K. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," IEEE Trans. Image Process., vol.14, no.11, pp.1918–1927, Nov. 2005.

[7] H.R. Sheikh, A.C. Bovik, and L. Cormack, "Blind quality assessment of JPEG2000 compressed images using natural scene statistics," Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers, vol.2, pp.1403–1407, Nov. 2003.

[8] H.R. Sheikh, Z. Wang, L.K. Cormack, and A.C. Bovik, "Blind quality assessment for JPEG2000 compressed images," Proc. Thirty-Sixth Annual Asilomar Conference on Signals, Systems, and Computers, vol.2, pp.1735–1739, Nov. 2002.

[9] H.R. Wu and M. Yuen, "A generalized block-edge impairment metric for video coding," IEEE Signal Process. Lett., vol.4, no.11, pp.317–320, Nov. 1997.

[10] L. Meesters and J.B. Martens, "A single-ended blockiness measure for JPEG-coded images," Signal Process., vol.82, pp.369–387, 2002.

[11] Z. Wang, A.C. Bovic, and B.L. Evans, "Blind measurement of blocking artifacts in images," Proc. IEEE Int. Conf. Image Processing, pp.981–984, Sept. 2000.

[12] K.T. Tan and M. Ghanbari, "Frequency domain measurement of blockiness in MPEG-2 coded video," Proc. IEEE Int. Conf. Image Processing, pp.977–980, Sept. 2000.

[13] A.C. Bovik and S. Liu, "DCT-domain blind measurement of locking artifacts in DCT-coded images," Proc. IEEE Int. Conf. Acoust. SPeech, and Signal Processing, pp.1725–1728, May 2001.

[14] D. Taubman and M. Marcellin, JPEG2000 Image Compression Fundamentals, Standards and Practice, Kluwer Academic Publishers, Boston, U.S.A., 2001.

[15] J.W. N1894, "JPEG2000 verification model 8.6 software, iso/iec," 2000.

[16] A. Bilgin, P. Sementilli, and M. Marellin, "Scalable image coding using reversible integer wavelet transform," IEEE Trans. Image Process., vol.9, no.11, pp.1972–1977, Nov. 2000.

[17] C. Perkins, RTP: Audio and Video for the Internet, Pearson Education, 2003.

[18] H. Cai and B. Zeng, "On packetization of JPEG 2000 code-streams in wireless chanels," IEEE International Conf. on Image Processing, vol.2, pp.841–844, Singapore, Oct. 2004.
[19] E. Edwards, "RTP payload format for JPEG 2000 video streams," Internet Draft, Oct. 2004.
[20] S. Futemma, A. Leung, and E. Itakura, "Rtp payload format for jpeg 2000 video streams," IETF Draft, http://www.ietf.org/internet-drafts/draft-ietf-avt-rtp-jpeg2000-10.txt, Feb. 2006.
[21] A. Leung, S. Futemma, and E. Itakura, "Payload format for jpeg 2000 video: Extensions for scalability and main header recovery," IETF Draft, http://www.ietf.org/internet-drafts/draft-ietf-avt-rtp-jpeg2000-beam-03.txt, Feb. 2006.

Kiyoshi Nishikawa received the B.E., the M.E., and the D.E. degrees in electrical engineering from Tokyo Metropolitan University in 1990, 1992 and 1996 respectively. From 1992 to 1993, he was at the Computer Systems Laboratory, Nippon Steel Corp. as a researcher. Since 1993, he has been with Tokyo Metropolitan University as an Assistant Professor and he is currently an Associate Professor of the Department of System Design. His research interest includes the adaptive signal processing, hardware architecture of digital signal processing, and digital communications. Dr. Kiyoshi Nishikawa is the secretary of the IEEE Signal Processing Society Japan Chapter and the Signal Processing Committee of the IEICE. He is a member of IEEE Signal Processing, and Communications.

Shinichi Nagawara received B.E. and M.E. degrees in electrical engineering from Tokyo Metropolitan University, Tokyo, Japan, in 2004 and 2006, respectively. He joined NEC Engineering, Ltd. in 2006. His research interests include image coding, streaming media and quality of service.

Hitoshi Kiya received his B.E. and M.E. degrees from Nagaoka University of Technology, Japan, in 1980 and 1982, respectively, and his D.E. degree from Tokyo Metropolitan University, Japan, in 1987. In 1982, he joined Tokyo Metropolitan University, where he is currently a Professor of Information and Communication Systems Engineering, Faculty of System Design. He was a Visiting Fellow at the University of Sydney, Australia, from Oct. 1995 to March 1996. His research interests include digital signal processing, multirate systems, image processing, and security for multimedia. Prof. Kiya has served as an Associate Editor for the IEICE Trans. Fundamentals (Japanese Edition) and the IEEE Trans. Signal Processing from 1998 to 2002 and from 1998 to 2000, respectively. He is a vice chair of the Signal Processing Committee of the IEICE and the chair of the Media Engineering Committee of the ITE, respectively. He is also a Member of the ITE and the IIEEJ, and a Senior Member of the IEEE.