An Adaptive Algorithm for the Quantization Step Size Control of MPEG-2

Nam-Ik Cho

Abstract

This paper proposes an adaptive algorithm for the quantization step size control of MPEG-2, using the information obtained from the previously encoded picture. Before quantizing the DCT coefficients, the properties of reconstruction error of each macro block (MB) is predicted from the previous frame. For the prediction of the error of current MB, a block with the size of MB in the previous frame are chosen by use of the motion vector. Since the original and reconstructed images of the previous frame are available in the encoder, we can calculate the reconstruction error of this block. This error is considered as the expected error of the current MB if it is quantized with the same step size and bit rate. Comparing the error of the MB with the average of overall MBs, if it is larger than the average, small step size is given for this MB, and vice versa. As a result, the error distribution of the MB is more concentrated to the average, giving low variance and improved image quality. Especially for the low bit application, the proposed algorithm gives much smaller error variance and higher PSNR compared to TM5 (test model 5).

I. Introduction

In MPEG-2 video encoding standard, picture blocks are transformed by DCT (discrete cosine transform) and the coefficients are adaptively quantized according to the given bit rate and picture properties[1, 2]. But detailed algorithms for each components of the MPEG-2 are not specifically recommended, giving flexibility in the design of the MPEG-2 encoder. In case of MPEG-2, the work group provides some test model of encoder where specific adaptive quantization scheme is also included[3, 4]. The TM5 (test model 5)[4] is considered as a standard MPEG-2 encoder and widely used for implementation and comparison. The adaptive quantization algorithm in TM5 is very simple, yet provides satisfactory result. However, TM5 considers only the current picture being encoded. In video coding system, we can obtain more information on the error distribution from the previously encoded pictures, which is available without additional frame memory. Thus, in this paper, we propose an adaptive quantization algorithm which exploits the error information obtained from the previously encoded pictures. The basic idea is to examine the error distribution of the macro blocks (MBs) of the previously encoded image and use this information for concentrating the error distribution of the currently encoded image. That is, the MBs with large error are more finely quantized by the proposed algorithm, resulting into the MB with smaller error. Conversely, the MBs with smaller error are quantized more coarsely. Since the visibility of artifacts is generally higher in the MBs with large reconstruction error, reducing large errors can enhance picture quality. Intensive computer simulation for the proposed algorithm is performed, and the result shows that the variance of the error distribution of the MBs is reduced and PSNR is also improved compared to TM5.

The overall structure of the encoder in this paper is the same as TM5, except that the adaptive quantization algorithm is replaced by the proposed one. Since the other video coding standards have similar adaptive quantization algorithms, the proposed one can be easily applied to other video coding standards.

This paper is organized as follows. In section II, the proposed algorithm is explained in detail. In section III, intensive simulation is performed for comparison with TM5. In section IV, we draw conclusions.

II. Adaptive Algorithm for Step Size Control

The adaptive algorithm for step size control in TM5 [4] is very simple and provides satisfactory result. The step size is adjusted according to the activity of the block and given bit rate. However, this approach use only the picture properties of the current frame for estimating the quantization error of the MB being encoded. In all video coding standards, the previous original and

Manuscript received September 15, 1997; accepted November 20, 1997.
N. I. Cho is with the School of Electrical Engineering, University of Seoul, Seoul, Korea.

reconstructed frames are always available without additional cost. This means we can get some information from these pictures for predicting the properties of quantization error of the current MB. When an MB is being encoded, if we can find similar parts of this MB from the previous frame, then we can predict the error properties of this MB because the original and reconstructed images of the previous frame are available. Finding the similar block in the previous frame is also already performed by the motion estimation component, and the result is given as a motion vector (MV). We can expect that the reconstruction error of this MB is similar to the difference between the motion compensated original and reconstructed block in the previous frame. But this does not mean that the magnitude of the error is exactly predicted with this simple scheme, because the bit allocator gives different number of bits to I, B and P images. However, the tendency of the error, whether the current MB will have larger or smaller error compared to other blocks or overall average, can be predicted. If the predicted error of the MB is larger than the average, the step size is reduced for fine quantization, and if it is smaller, coarse quantization is performed by increasing the step size. As a result, the distribution of the reconstruction error of each MB will be more concentrated to the average. That is, the formerly large errors will be decreased and small errors will be increased. Since the visibility of the artifacts is higher in the MBs with higher reconstruction error, it is believed that this result will give better image quality. Intensive simulation results will be give in the next section, for demonstrating that the error variance is reduced and in many cases the PSNR is also improved.

Now, let us describe the proposed algorithm more precisely. In case of MPEG-2, the images are entitled as I, B, or P and displayed in the order as shown in Fig. 1. However, they are actually encoded in the order of I, P, B, B, P, B, B, P, $\cdots$. In coding the P image, the previous image is either I or P. So we should find the predicted (motion estimated) block with forward MV. In case of B, the forward and backward MVs are available. Either of these vectors or both are used for the motion compensation of the MB. However, intensive simulation shows that using the nearest vector gives slightly better result for the proposed algorithm. That is, forward vector is used for the B frame that comes first, and only backward vector is used for the second B frame.

In summary, for finding the similar part of the current MB from the previously encoded frame, forward or backward MVs are used. In case of I, the first I is encoded the same as the TM5. The rest of I frames are encoded using previous P frame. However, since there is no MV available for I images, the MV is set to (0,0). In other words, the MB in the same position of previous P frame is referenced for predicting the MB reconstruction error.

The overall encoder structure is the same as TM5. The quantization is processed as follows.

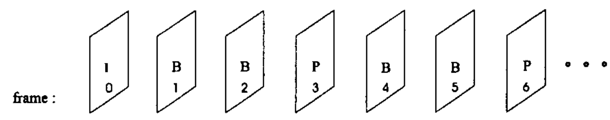


Fig. 1. Display order of images in MPEG-2.

1. The first I frame is encoded the same as the TM5.

2. After encoding I or P image, the difference image is made using the available original and reconstructed frames. This is used for calculating the MB reconstruction error for the next frames in the form of SAD (sum of absolute difference) or other errors. For saving the frame memory, difference image may not be made, but SAD for MB is computed by using original and reconstructed frame every time when an MB is referenced.

3. In case of encoding P, the forward MV is used for pointing the most similar parts of the current MB. A block with the size of MB is picked in th [illegible] D is computed. In case of B, either of the forward or backward MV can be used. But the nearest one is used for computing SAD. The SAD is used as the error measure in this paper, however, various kinds of error can be considered. In case of I, that is not the first, the MV is given (0,0) and the SAD for each MB is computed with reference to the previous P frame. Let us define an array that contains the SAD of each MB be earray[]. Of course, the size of the array is the number of MBs in a frame, and earray[k] has the predicted SAD k-th MB.

4. Denoting the number of quantization steps as mquant, the main part of the adaptive quantization algorithm is performed as follows.

```
for(j=0; j<B; j++) {
Compute mquant as in TM5;
ratio = earray[j]/average;
mquant = mquant/ratio;
Clipping and refinement of mquant as in TM5;
}
```

where B is the number of MBs in a frame, and

$$\text{average} = \frac{1}{\mathrm{B}} \sum_{k=0}^{\mathrm{B}} \text{earray}[k]. \tag{1}$$

In the above algorithm, ratio is greater than 1 if the current error of MB is larger than the average. Hence, the mquant produced by TM5 algorithm becomes smaller by dividing it with ratio. In MPEG-2 standard, mquant can be adjusted from 1 to 31, and larger mquant means coarse quantization[4]. Conversely, the MBs with smaller error than the average, the **ratio** is less than 1 and they are coarsely quantized than before.

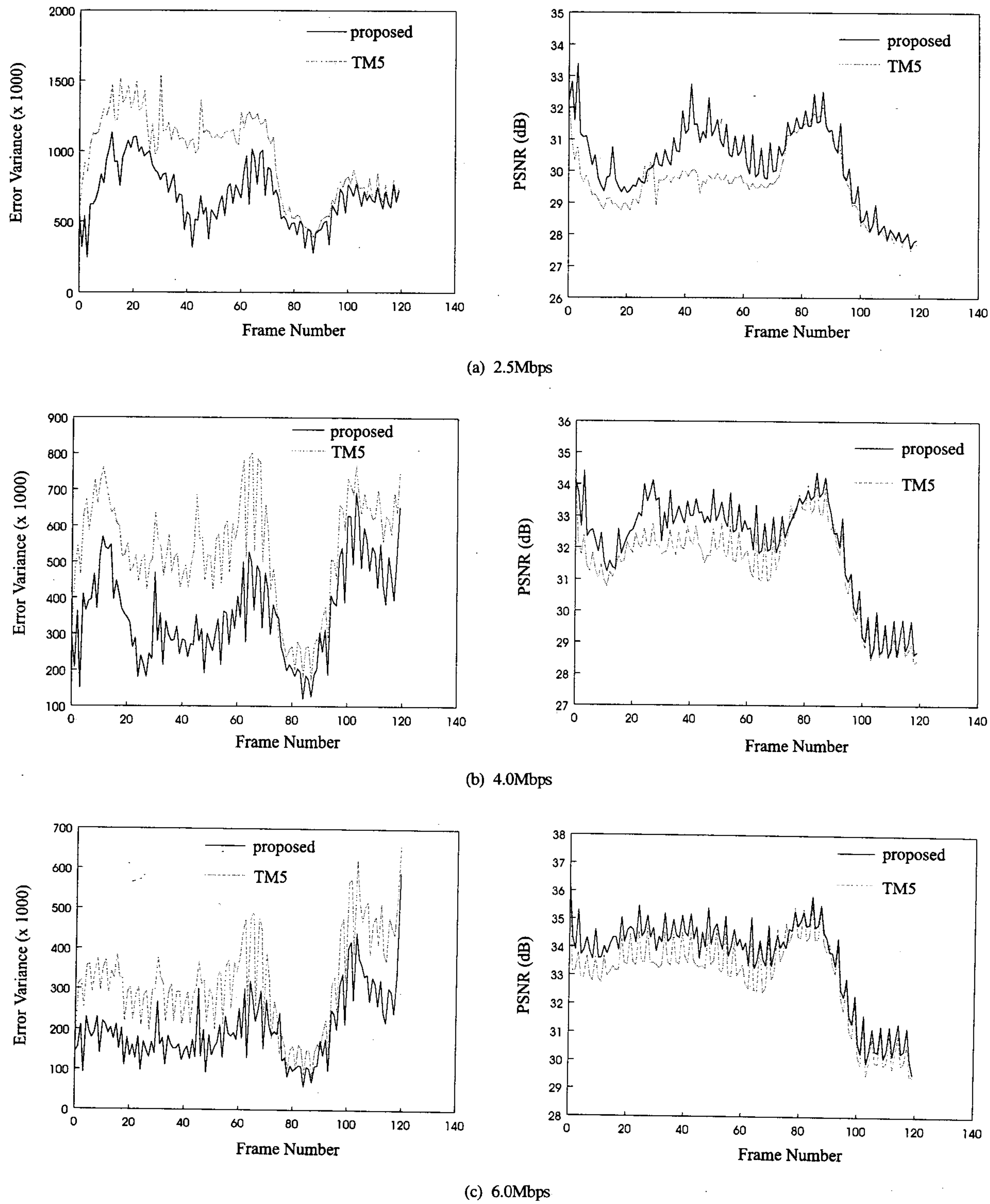


Fig. 2. Error variance and PSNR of popple sequence.

III. Comparison with TM5

In order to demonstrate that the proposed algorithm reduces the error variance, several image sequence with different properties are encoded in various bit rates. The "variance" in this paper is the variance of the reconstruction errors (SAD) of each MB, which is defined by

$$\text{variance} = \frac{1}{B} \sum_{k=0}^{B} (\text{earray}[\,k\,] - \text{average})^2 \qquad (2)$$

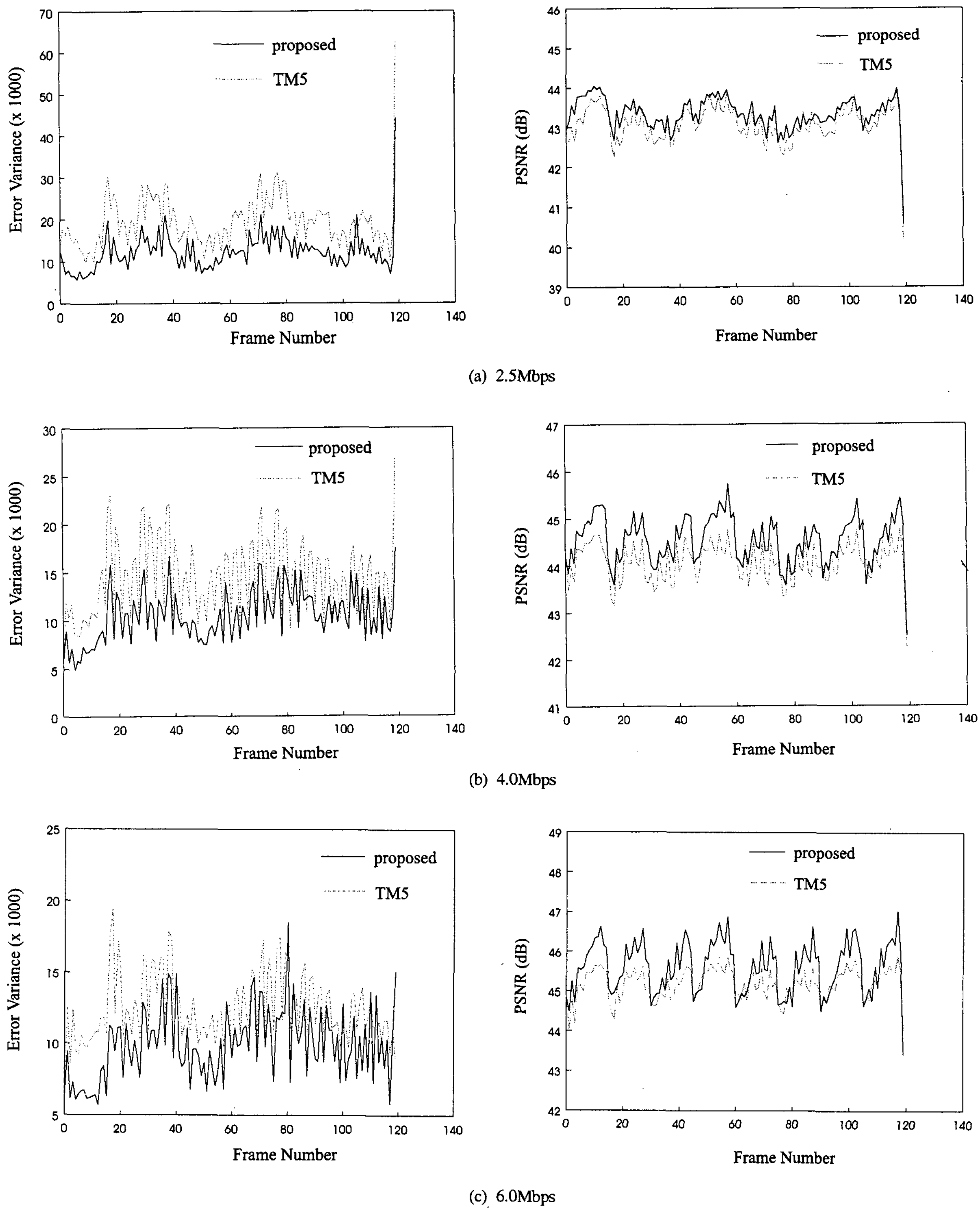
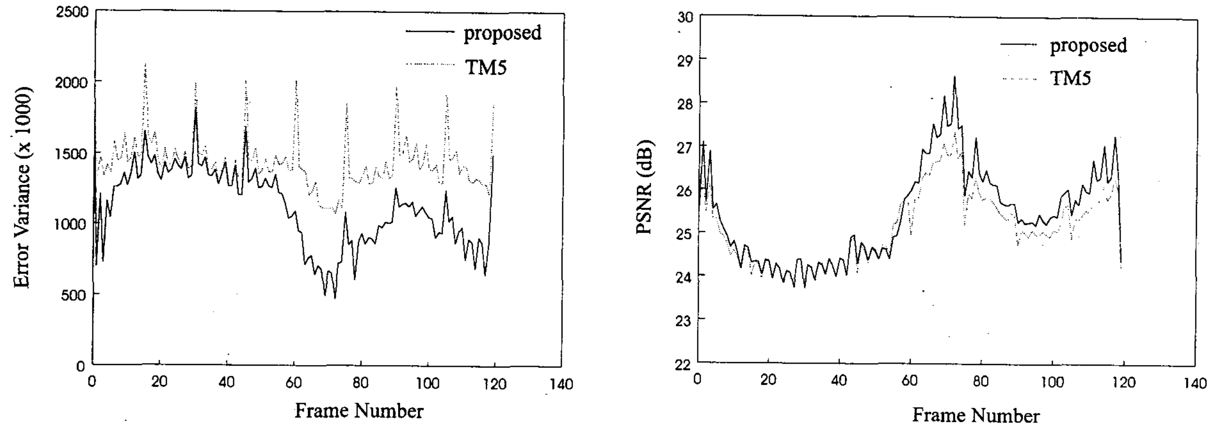


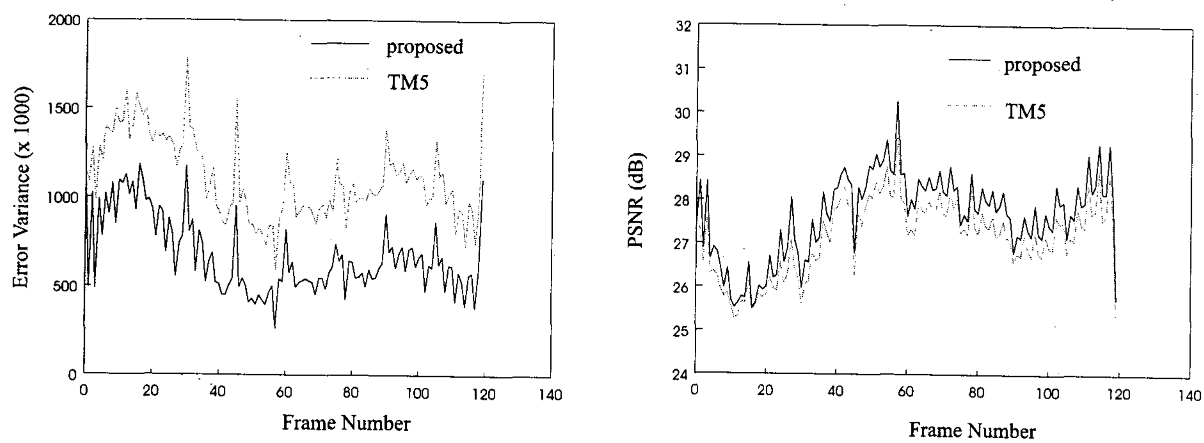
Fig. 3. Error variance and PSNR of akiyo sequence.

where each of the variables are defined the same as equation (1). Of course, the smaller variance means that the reconstruction errors of MBs are not much different from each other. For detailed comparison of the algorithm with the TM5, intensive simulations have been performed. For each of the image sequences, popple, akiyo, mobile and calendar, and football, proposed algorithm and TM5 are performed with the bit rate of 2.5, 4, and 6Mbps. From Fig. 2 to Fig. 5, error variance and PSNR of the proposed algorithm and TM5 are compared for each case.

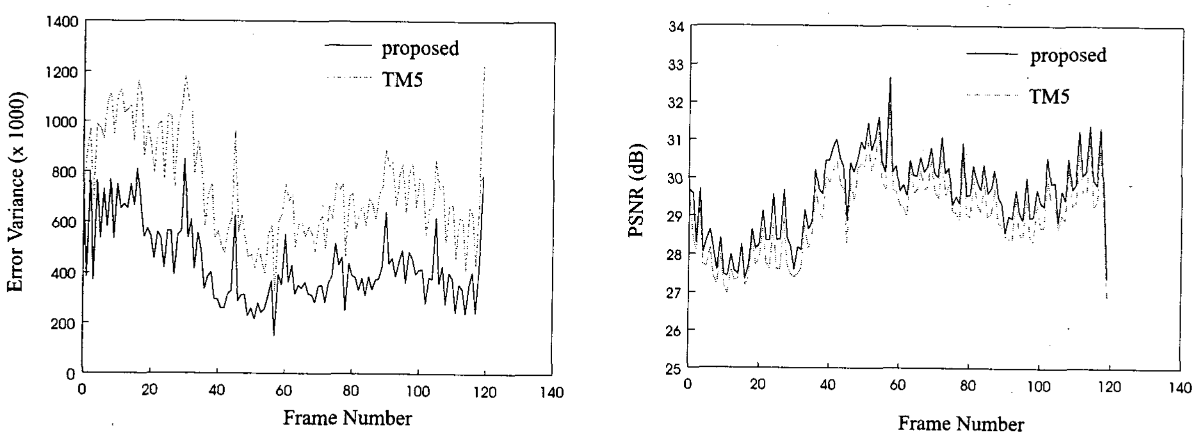
From the figures, it is believed that the proposed algorithm



(a) 2.5Mbps



(b) 4.0Mbps



(c) 6.0Mbps

Fig. 4. Error variance and PSNR of mobile and calendar sequence.

improves the quality of the reconstructed image compared to TM5, especially for the low bit rate application and slow motion pictures. The point of the result is not that the PSNR is improved, but the error variance is reduced, because the algorithm is intended to reduce the error of the MB which has larger error than the average, rather than to reduce the overall reconstruction error. By reducing the error variance, the MBs with very large error now have smaller error. This can also be shown more visually as in Fig. 6, which are the reconstructed and error images produced by TM5 and the proposed algorithm. The error images are not difference image, but SAD is represented as brightness for each MB. The brighter block in Fig. 6 is the MB with higher

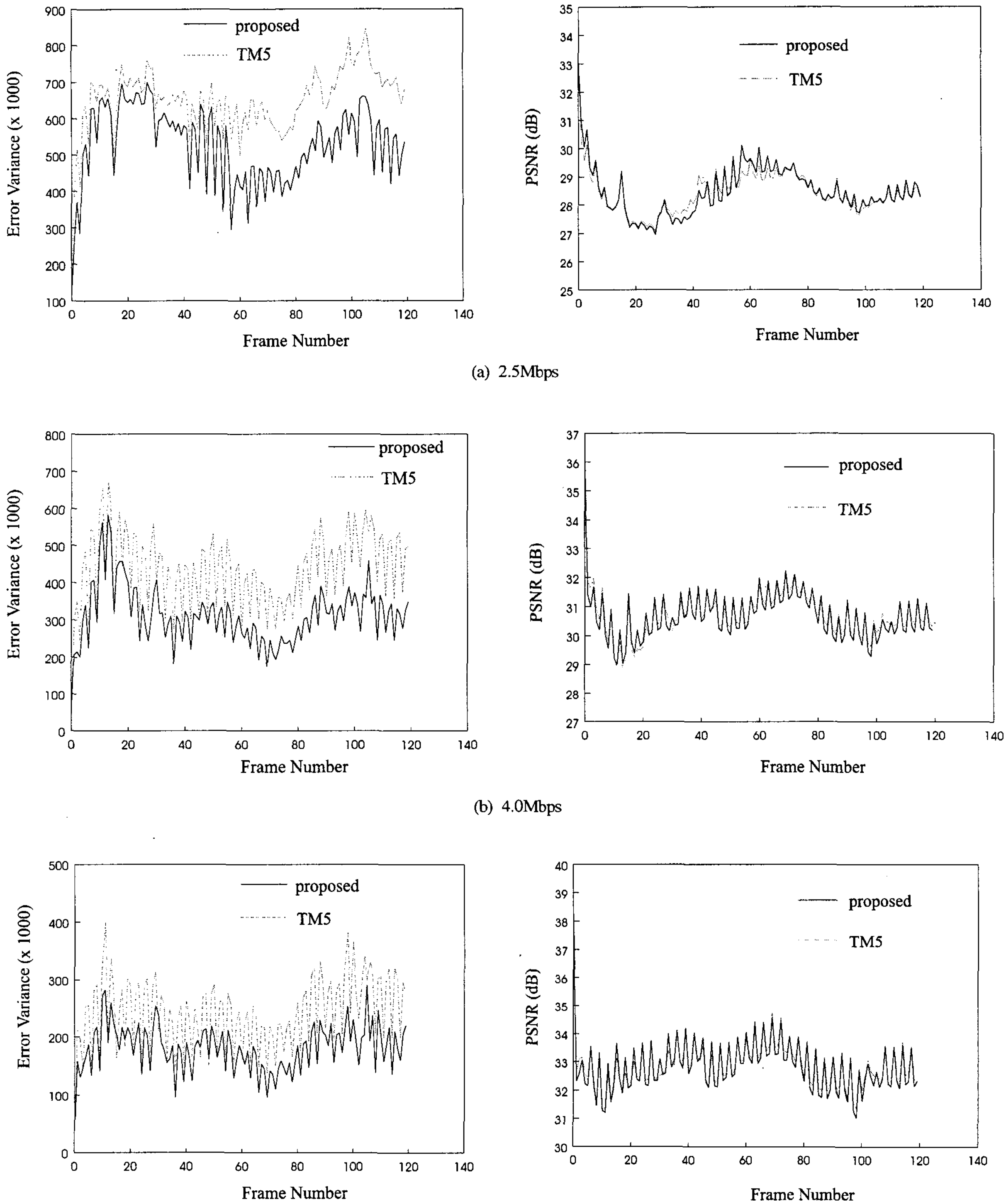
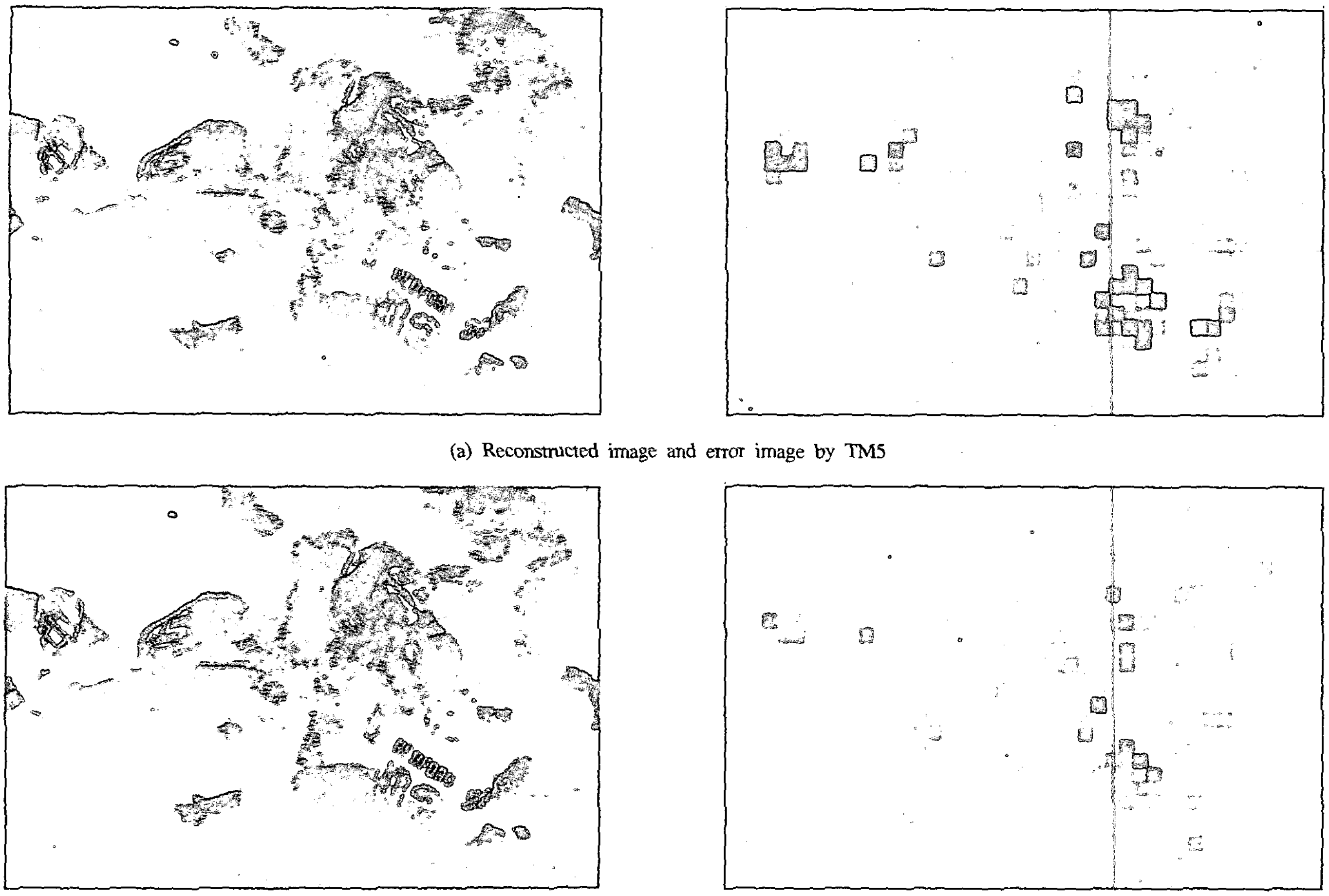


Fig. 5. Error variance and PSNR of football sequence.

reconstruction error. The figure shows that very bright blocks in Fig. 6(a) with higher SAD became less bright in the image produced by the proposed algorithm as in Fig. 6(b). In general, the MB with larger SAD has higher visibility of artifacts. Hence, by reducing the larger errors, the overall picture has less visible artifacts. However, since there are many additional factors that affects the visibility of the artifact such as brightness, frequency, edge, motion, and etc[5-8], the error measure based on human visual perception should also be considered rather than reducing SAD only, which is left as a future research topic.



(a) Reconstructed image and error image by TM5

(b) Reconstructed image and error image by the proposed algorithm.

Fig. 6. Reconstruction and error image by TM5 and the proposed algorithm(20th frame).

IV. Conclusions

We have proposed an adaptive quantization algorithm for video coding. The step size of the quantizer is adjusted using the information obtained from the previously encoded image, whereas only the picture property of the current image is used in the existing algorithms. Since all the information used in the proposed algorithm is already available in the video encoding standards, the implementation of the algorithm is very simple. From the intensive computer simulation, it is shown that the error variance of the proposed algorithm is smaller than that of the TM5 in almost all cases, and PSNR is also improved in many cases especially for the sequences with slower motion and for low bit rate encoding. This means that the MBs with very large reconstruction error turned to have smaller error by the proposed algorithm. We have also proposed a modified algorithm for more efficient hardware implementation.

References

[1] Motion Picture Expert Group, MPEG Draft International Standard, Technical Report, CCITT SGXV, March 1994.

[2] D. Le Gall, "MPEG : A video compression standard for multimedia applications," *Commun. ACM*, Vol. 34, pp. 47-58, April 1991.

[3] Motion Picture Experts Group, "Description of simulation model 3, MPEG 90/041, July 1990.

[4] ISO/JTC1/SC29/WG11/N0400, "Coded representation of picture and audio information-MPEG2 Test Model 5," April 1993.

[5] S. A. Karunasekera, N. G. Kingsbury, "A distortion measure for blocking artifacts in images based on human visual sensitivity," *IEEE Trans. on Image Processing*, Vol. 4, No. 6, pp. 713-724, June 1995.

[6] K. N. Ngan, "Hybrid image coding scheme incorporating human visual system characteristics," *Optical Engineering*, Vol. 30, pp. 940-946, July 1991.

[7] N. Li, S. Desmet, A. Deknuydt, L. V. Eycken, "Motion adaptive quantization in transform coding for exploiting motion

masking effect," *SPIE Vol. 1818 Visual Communication and Image Processing '92*, pp. 1116-1123, 1992.

[8] S. K. Paek, J. S. Kang, Y. S. Seo, "Rate control strategy based on human visual se-nsitivity for MPEG video coder," *SPIE Vol. 2308 Visual Communication and Image Processing '94*, pp. 322-330, 1994.

Nam-Ik Cho received the B.S., M.S. and Ph.D. degree from Seoul National University, Seoul, Korea, in 1986, 1988 and 1992, respectively, in control and instrumentation engineering. He is now with the School of Electrical Engineering, University of Seoul, Korea. His research interest is in digital signal processing, including adaptive filtering and VLSI implementation.