An Efficient Requantization for Transcoding of MPEG Video

Hee-Chul Hwang, and Duk-Gyoo Kim
School of Electronic and Electrical Engineering, Kyungpook National University
1370, Sankyug-Dong, Buk-Ku, Daegu 702-701, Korea
Tel: +82-53-940-8625, Fax: +82-53-950-5505
E-mail: heeyaa@palgong.knu.ac.kr

Abstract : In this paper, we propose an efficient transcoding of MPEG video. Transcoding is the process of converting a compressed video format to another different compressed video format. We propose an simple and efficient transcoding by requantization in which MPEG coded video at high bit-rate is converted into MPEG bitstream at lower bit-rate. To reduce a image quality degradation, we use HVS(Human Visual System) that is the effect that visibility of noise is less in high activity regions than in low activity regions. By using the effect, the part of image in high activity region is coarsely quantized without seriously degrading the image quality. Experimental results show that the proposed method can provide good performance.

1. Introduction

MPEG(Moving Picture Experts Group) is an efficient video and audio coding standard. And it is used to many video transmission and storage applications[1]. Specially MPEG-1 is used for Video CD and MPEG-2 video is used in the digital broadcasting , DVD applications. MPEG-4 is for a broadband network such as internet applications or mobile/wireless network. In the storage applications, high quality video can be stored in DV or MPEG-2 format or DVD applications, whereas MPEG-1 and MPEG-4 would be preferred for video mail, video delivery and CD-ROM application purposes. In many non-realtime applications, a video source is compressed at a predetermined quality or bit rate. The compressed video is stored for storage or transmission at a late time. In order to transmit a compressed video signal over a network with limited bandwidth, the compressed bit stream needs to be decoded and re-encoded, or transcoded for the desired channel rate.

For example, a program provider transmits a digital video signal across a satellite link in a compressed format and this program is to be relayed on a cable network. The transmitted signal must be in the more compressed format.

In this way, various video transcoding techniques have been developed to transmit or store multimedia contents efficiently in such network or storage applications by converting compressed bit streams into lower rates recently. A straightforward approach is to first fully decode the compressed video and then re-encode it at a lower bit-rate.

Transcoding is the process of converting between different compression formats or reducing the bit rate of a previously compressed signal. Fig.1. is a basic block diagram of a video coding system.

Video transcoding discussed in [2] provides several main concepts of MPEEG-2 bit rate scaling. It is very

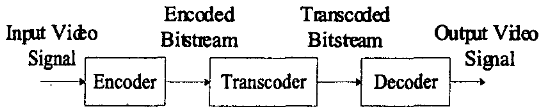

Figure 1. A video coding system including a trancoder

time-consuming task to encode the original video signal to each target format.

We propose an efficient trasncoding algorithm of MPEG-2 to other format with the reduced bit-rate. An important issue in bit rate conversion is to provide requantization methods for efficient transcoding[3,4].

Basically, function of transcoding can be realized by connection decoder and encoder for appropriate format. This is the easiest way to transcode coded video date into another coding format, but this model may cause quality-loss of images.

In this paper, we propose quality-effective video transcoder. Proposed transcoder uses HVS(Human Visual System) in order to reduce image quality degradation. Proposed transcoder uses quantizer parameter, macroblock coding type, DCT coefficients and prediction errors.

To reduce image quality degradation, a block having edges can be found from the DCT coefficient of the block and by detecting prediction errors from difference images between the original and the compensated images, it is able to re-decide the quantization step size. Not only using the traditional method of MPEG but also using these informations, the adaptive requantization is performed for compensating the error from quantization.

2. MPEG Transcoding

Video sequences usually contain statistical redundancies in both temporal and spatial directions and MPEG compress video data efficiently by reducing redundancies existing in an image sequence[5][6].

The basic statistical property is inter pixel correlation. So many video compression algorithms uses the assumption that the magnitude of a particular image pixel can be predicted from nearby pixels within the same frame or from pixels of a nearby frame. The MPEG compression algorithms employ DCT coding techniques in image blocks of 8 x 8 pixels to explore spatial correlations between nearby pixels within the same image. The term intra is performed relative to information that is contained only within the current frame, and not relative to any other frame in the video sequence. However, if the correlation between pixels in nearby frames is high, it is desirable to use an inter-frame coding techniques employing temporal prediction.

In MPEG video coding schemes an adaptive combinations of both temporal motion compensated prediction followed by transform coding of the remaining spatial information is used to achieve high data compression.

The algorithms of MPEG are the hybrid of motion estimation, DCT transform, quantization. The standard specifies only the decoding algorithm and bit-stream syntax in detail. It is thus possible to have different encoders for optimal trade-off of complexity and performance. As stated above, an important feature of the MPEG encoding algorithm is the flexibility of the bit-rate to specific applications requirements by adjusting the quantization step size to quantize the DCT-coefficients. Coarse quantization of the DCT-coefficients achieves high compression ratio in the storage or transmission of video, however, it may result in significant coding artifacts. The MPEG standard allows the encoder to select different quantizer values for each coded macroblock. This enables a high degree of flexibility to allocate bits in images where needed to improve image quality.

In current video encoders, a motion-compensated predictive coding scheme is used to reduce the temporal redundancy between consecutive frames. In this scheme, the motion-compensated previously reconstructed picture is subtracted from current frame, and the difference error signal is encoded and transmitted to the decoder. To reconstruct the frame correctly in the decoder, the encoder reference should be exactly the same as that reconstructed by the decoder. Otherwise, the mismatched reference picture will produce the distored reconstructed image in the decoder. Since the distored reconstructed image is also used for the future prediction, the distortion error propagates to the future reconstructed pictures. Therefore, even a small mismatch at one frame can cause significant quality degradation over time.

A *transcoder* located between a front-encoder and an end-decoder is shown in Fig. 2.

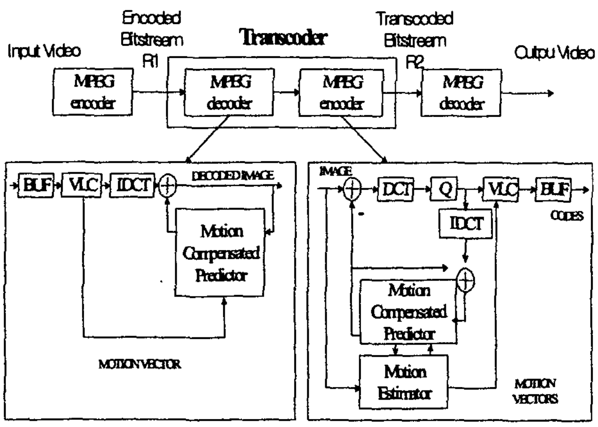

Figure 2. A transcoder in MPEG video codec

This figure shows a basic block diagram of a video coding and decoding system that includes a transcoder. An encoder compresses the input video signal at a bit rate of R1 Mbit/s, then the transcoder converts this compressed signal into a compressed format of a lower bit rate R2 Mbit/s and next a decoder decompresses the incoming transcoded video signal and finally display the resulting video signal.

In MPEG standards, the quantization step size includes two components, a quantization coefficient specifying a minimum step size for the particular DCT coefficient, and a quantization scaling parameter, that is called by MQUANT, for bit-rate control. In general, the bit-rate reduction of compress video can be achieved by either coarse quantization, or by discarding high frequency coefficients.

In this paper, transcoding is achieved by requantization of the DCT coefficients.

3. Adaptive Transcoding

Transcoding by means of cascading a decoder and an encoder may lead to several types of problems such as complexity and performance at a given bit rate. We focus on reducing the image quality degradation.

Video with different feature may not necessary have the same perceptual quality when they are coded at the same bit-rate. For example, to attain the similar perceptual quality, video with rapid scene changes requires more bits to code than video with mainly stationary scene. Video with many spatial details demands more bit than video with uniform scene. The quality of a compressed video can often be traded for the desired video bit-rate. To get a better image quality, we use HVS(Human Visual System) that is the effect that visibility of noise for human visual system is less in high activity regions than in low activity regions. By using the effect, the part of image in high activity regions is coarsely quantized without seriously degrading the image quality. But, in this way, the reconstructed images have the degradations in the blocks having edges or fast and complex motion that is not well compensated normally. By using DCT coefficients and difference image as well as image activity, the degradations of image quality in those area can be reduced.

As for quantization, activity of image is most essential parameter in pixel domain to analyze the complexity of image characteristics such as texture, edge, and flat region. It is used to determine quantization step and bit allocation. For example, in MPEG-2 TM(TestModel) 5, minimum block variance in macroblock is used as local activity in the picture. Here is how we get the activity measurement, that is mquant, from image.

To get a mquant, first step is to compute a spatial activity measure for the macroblock j from the four luminance frame-organised blocks(n=1...4) and the four luminance field-organised blocks (n=5...8) using the intra pixel values:

$$act_j = 1 + \min(vblk_1, vblk_2, \ldots, vblk_8)$$

where,

$$vblk_n = \frac{1}{64} \times \sum_{k=1}^{64} (P_k^n - P\_mean_n)^2 \quad \text{and}$$

$$P\_mean_n = \frac{1}{64} \times \sum_{k=1}^{64} P_k^n$$

where avg_act is the average value of *actj* the last picture to be encoded. Then *mquantj* is obtained as,

$$mquant_j = Q_j \times act_j$$

where *Qj* is the reference quantization parameter obtained by rate control. Consequently mquant is the quantization parameter considering the visual importance and deciding the quantization step. Quantization is done by the multiplication of mquant and quantization matrix. We use the information of DCT coefficient and prediction error as well as the activity of a image. It is for considering edge regions and prediction errors. The proposed method is based on the adaptive quantization and rate control for requantization by reflecting HVS. It uses the effect that visibility of noise is less in regions of high activity than in regions of low activity. By using the effect in regions of high activity the image is coarsely quantized without seriously degrading the image quality.

(1) Intra macroblocks
In the process of DCT transform, we pick up two sets of adjacent pixels out of reconstructed above and left block illustrated by Fig.3

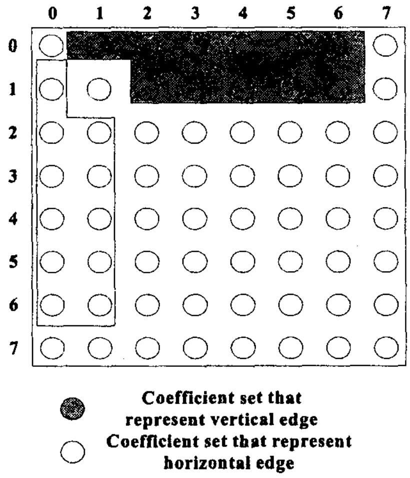

Fig 3. DCT coefficients sets

AC coefficients bearing edge regions of current block are related to the ACs of adjacent DCT sets. Vertical edges are shown in a block if the magnitude of above set is high, and horizontal edges are shown if left set is high respectively. The steep of edge can be determined from the ratio of above and left AC magnitudes. As these ratios are reflected to the mquant, quantization noise came from coarse quantization on expected edge can be prevented. If ratio is high, it means there are steep edges, mquant become smaller. So the finer quantization is obtained.

(2) Inter macroblocks
The regions with low spatial activities are already quantized finely in MPEG TM5. We consider that the fast and complicated blocks not well compensated. After motion prediction, the difference image is achieved from the subtraction of original and prediction image. By summarizing all prediction errors for all blocks, average prediction error per a frame can be calculated. And a error in a block can be shown as ratio of a error in a block and a average error in a frame. If the ratio is high, it means there are blocks that are not well compensated and mquant become smaller. So the finer quantization is done. Consequently mquant is recontrolled by depending on intra or inter MB type. For intra blocks, the information of edge is included in mquant by DCT coefficients. For inter blocks, the information of prediction error is included in mquant by difference images.

```
for(j=0; j<Block; j++)
{
  Compute mquant as in TM5;
  if (MB=Intra)
    ratio= vertical coefficients of DCT
          / horizontal coefficients of DCT
      if (ratio <1) ratio=1/ratio
      else ratio=ratio
  else
    ratio=error[k] of a MB/average_error _image;

  mquant_final=mquant/ratio
}
```

where mquant and $mquant_{final}$ are quantization step before and for quantization at macroblock j in k th frame respectively. In this model, adaptive quantization is performed through analysis of activity and the quantization step is obtained by multiplying base quantization step by a weighting factor determined by activity that is including the information of edge parts and prediction errors. The base quantization step is given by bit allocation and buffer control process among macroblocks, picture, and GOP(Group of Picture) layers. Since lower activity region is to be perceived more sensitively from HVS point of view and to generate less bit counts than that of higher activity region at the same quantization step, smaller quantization step is used for lower activity region. Therefore, the weighting factor is given as the ratio depending on intra or inter macroblock.

4. Simulation Results

The performance of the proposed scheme is investigated by computer simulation in comparison with the encoding only scheme in the MPEG-2 TM5. As test images, various types are used for an objectivity; 100 frames of 352×240 size Football, Flower Garden, Mobile, Table Tennis that is SIF (source input format) sequence. Football has somewhat large and irregular motions, Flower Garden has wholely regular motions by camera panning and complicated edge regions, Mobile and Table Tennis have partial motions and scene changes. The sequence is first encoded with an MPEG-2 TM5 encoder and then transcoded at different bit-rate. Commonly, the comparisons are carried out by objective and subjective image quality and the amount of generated bits. The most commonly used image quality criterion, PSNR (peak signal to noise ratio), is used to compare objective quality.

Since the results obtained are similar for all sequences, so here only Flower Garden sequence is used to show performance of the proposed method. The image quality obtained by transcoding a 4 Mbit/s bit stream of Flower Garden sequence into 1.5Mbit/s using the proposed algorithm is compared with that of the corresponding encoded only at the same rate of 1.5Mbit/s in figure.4.

From this figure, it can be seen that the proposed algorithm results in a image quality have similar performance to that of images by the method of encoding only. The average PSNR of the other sequences like Football, Mobile and Table Tennis are compared in Table 1 for the comparison of performance. As the table shows, the results of the proposed method are very close to encoding only method. The PSNR of Flower Garden sequence with regular motions has the best performance as the difference is 0.11dB and other sequences have PSNR between 0.1 to 0.3 dB.

However, coding SIF pictures at 4 Mbit/s results in very high quality, although the input of the transcoder is a compressed bit stream, so the distortion before transcoding is small. The difference would be greater than for a lower bit-rate transcoding. Further research work will be carried on this part.

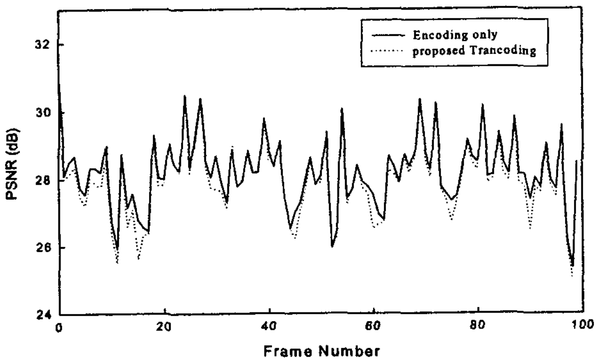

Fig 4. PSNR comparison of Flower Garden

Sequence	Encoding only (dB)	Proposed method(dB)	Difference
Football	29.25	28.93	0.32
FlowerGarden	28.19	28.08	0.11
Mobile	26.78	26.60	0.18
Table Tennis	34.25	34.01	0.24

Table 1. Comparison of the proposed method performance

5. Conclusion

In this paper, we propose effective methods that can reduce the degradations of image quality comes from requantization. Our main concern is in providing a simple yet efficient way to reselect the most suitable quantization step size according to image feature. By using the proposed method, the transcoding efficiency is similar to that of encoding only. Quantization parameter in the traditional quantization of MPEG only uses the activity of image but in the proposed paper, quantization parameters are recontrolled according to the image feature with the information from DCT coefficient and difference image for considering image quality. So the block needed the finer quantization can be quantized finely for better image quality. Experimental results shows that the proposed method has better performance comparing to traditional quantization and has very close performance to encoding only method.

References

[1] D. Le Gall "MPEG: A video compression standard for multimedia applications," Commun,ACM,vol. 34, pp.47-58,Apr. 1991

[2] W.K.Sun, J.W.Zdepski, Architectures for MPEG com pressed bitstream scaling," *IEEE Trans. Circuit SystemVideo Tech*, vol. 6, no.2, pp. 191~199, 1996

[3] O.Werner, "Requantization for Transcoding of MPEG-2 Intraframes," *IEEE International Symposium on Circuits and Systems(ISCAS'99)*, vol. 8, pp. 179~191, Feb 1999.

[4] P.Assuncao and M.Ghanbari, "A Frequency Domain Video Transcoder for Dynamic Bit-Rate Reduction of MPEG-2 BitStreams," *IEEE Trans. on Circuits and Systems for VideoTechnology*, vol. 8, pp. 953~967, December, 1998.

[5] K.R.Rao and J.J.Hwang: Techniques and Standards for Image, Video and Audio Coding, Upper Saddle River, NJ: Prentice Hall, 1996

[6] J.L. Mitchell, W.B. Pennebaker, C.E.Fogg and D.J. Legall: MPEG Video Compression Standard, Champman and Hall, 1997