논문 01-02-S-1

An Efficient Requantization Method for INTRA Frames in Heterogeneous Transcoding

이종의 영상부호화 표준간의 변환부호화에서 화면내 부호화를 위한 효율적인 재양자화 기법

Kwang-Deok Seo*, Jae-Kyoon Kim*
(徐 光 德* , 金 在 均*)

Abstract

In this paper, we propose an efficient requantization method for INTRA frames in heterogeneous transcoding from MPEG-1 to MPEG-4 simple profile. The quantizer for MPEG-1 INTRA MB usually uses a quantization weighting matrix while the quantizer for MPEG-4 simple profile doesn't. As a result, the quantization step sizes of the two quantizers may not be the same even for the same quantization parameter. Due to this mismatch in the quantization step size, the transcoded MPEG-4 sequence suffers from serious quality degradation and the number of bits produced by transcoding increases from the original MPEG-1 video sequence. To solve these problems, we propose an efficient method to find a near-optimum reconstruction level in the transcoder. We also present a PDF (probability distribution function) estimation method for the original DCT coefficients of MPEG-1 video sequence, which is required for the proposed requantization. Experimental results show that the proposed method gives 0.3~0.6dB improvement in PSNR over the conventional method, even at the reduced bit-rate about 5~7% from the conventional method.

요 약

본 논문에서는 MPEG-1을 MPEG-4 심플 프로파일로 변환 부호화할 때 화면내 부호화를 위한 효율적인 재양자화 기법에 대해 제안한다. MPEG-1의 화면내 부호화 블록의 양자화는 양자화 가중 행렬을 사용하는 반면, MPEG-4 심플 프로파일은 양자화 가중 행렬을 사용하지 않는다. 그 결과 두 부호화 방식의 양자화에 사용되는 양자화 파라미터가 동일하더라도 양자화 계단 크기가 서로 달라지기 때문에 변환 부호화된 MPEG-4 영상의 화질이 심하게 열화 된다. 이 문제를 해결하기 위해 변환 부호기에서 양자화 오차를 최소화하는 재생레벨을 결정하는 방식을 제안하며, 이 방식의 적용을 위해 변환부호기에서 MPEG-1 시퀀스의 DCT 계수에 대한 확률밀도함수를 추정하는 방법을 제시한다. 실험결과에 의하면 제안된 방식을 적용할 경우 기존의 방식에 비해 PSNR 측면에서 0.3~0.6dB 정도의 개선이 있으며, 동시에 발생 비트량을 5~7% 정도 줄일 수 있다.

I. Introduction

Transcoding involves the sequential processing of partial decoding of a bitstream, possibly modifying the

---

* 韓國科學技術院 電子電算學科
(Dept. of Electrical Engineering and Computer Science, KAIST)

修正完了日: 2001年11月 9日

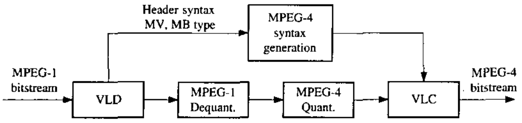


Fig. 1. Outline of the heterogeneous transcoder.

decoded contents, and re-encoding [7-13]. The transcoding is useful in a variety of applications. The transmission of compressed video over channels with lower bandwidth may require a reduction in bit rate through transcoding [7,10,11,12,13]. For other applications, one may need to produce in real-time a compressed bit-stream containing the video at a lower resolution than the original compressed video [7,9]. Note that the manipulations have been generally performed between the same video standard, which we call a homogeneous transcoding. However, in our work, video transcoding is regarded as a heterogeneous process of converting an MPEG-1 video sequence into corresponding MPEG-4 sequence. Only a few works have previously dealt with heterogeneous transcoding problem [8,9]. Acharya and Smith have studied heterogeneous transcoding from MPEG-1 to MJPEG in [8]. Shanableh and Ghanbari have done work on heterogeneous transcoding from MPEG-1,2 to H.263 [9]. They focused on determining new motion vectors for lower spatial resolution pictures and reformatting the incoming bitstreams sequence structure by employing PB-frame which is an optional mode of H.263+. In this paper, we consider the requantization problem in the transcoding from MPEG-1 video sequence into corresponding MPEG-4 simple profile sequence.

Fig. 1 shows the outline of heterogeneous transcoding from MPEG-1 bitstream to MPEG-4 simple profile bitstream. The input MPEG-1 bitstream is decoded by the variable length decoder (VLD), which gives the values of the quantized DCT coefficients in addition to various coding information, such as quantization parameter, MB type and motion vector. The quantized DCT coefficients are dequantized and then requantized with MPEG-4 quantizer. The requantized coefficients are encoded by the variable length coder (VLC) of MPEG-4 along with other information in conformance to MPEG-4 syntax. Fig. 2 shows the conversion flows of the important syntax of MPEG-1 to its counterpart of MPEG-4 in a hierarchical coding structure.

The main problems in heterogeneous transcoding are the syntax mismatch and any difference in coding algorithms between the associated video standards. In our work, the main differenct coding block is the quantizer. Usually the quantizer for MPEG-1 INTRA MB uses a quantization matrix while the quantizer for MPEG-4 simple profile doesnt. As a result, the quantization step sizes of the two quantizers may not be the same even for the same quantization parameter. Due to this mismatch in the quantization step size, the transcoded MPEG-4 sequence suffers from serious quality degradation and excessive bit generation. By analyzing the quantizer characteristics of MPEG-1 and MPEG-4, we propose an efficient requantization method to find, in the sense of MSE, a near-optimum reconstruction level in the transcoder. For the proposed method, we also present a PDF (probability distribution

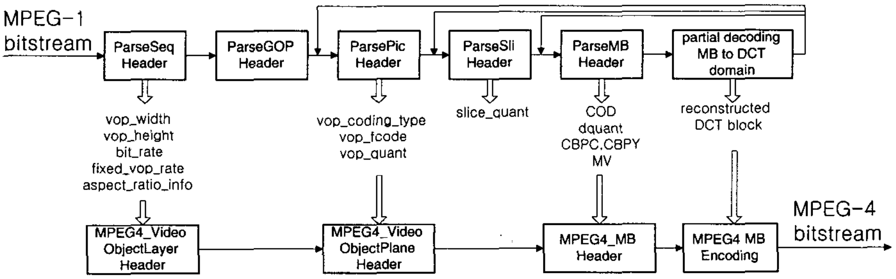


Fig. 2. Conversion flow of the important syntax.

function) estimation method for the original DCT coefficients of MPEG-1 video sequence in the transcoder.

The organization of this paper is as follows. In section II, we derive the quantizer characteristics for INTRA MB of MPEG-1 and MPEG-4. In section III, an efficient requantization method is proposed, and a simple method for estimating the statistics of the original DCT coefficients is presented in section IV. Section V shows some experimental results and the conclusion is provided in section VI.

II. Quantizer characteristics for INTRA MB OF MPEG-1 and MPEG-4

2.1 Quantization Characteristics for INTRA MB of MPEG-1

As the DC coefficient is separately handled for both codecs, we only pay attention to AC coefficients. A description of the quantizer characteristics of MPEG-1 INTRA MB is given in Fig. 3-(a). Fig. 3-(b) shows the default quantization weighting matrix for INTRA MB, where the weighting factor wi depends on the frequency index i, i=1, 2, .· , 63. With this weighting factor, the quantization step size Δ is specified as[3]

$$\Delta = \frac{w_i \cdot Q_p}{8} \qquad (1)$$

where Qp is the quantization parameter with 1,2,... ,31.

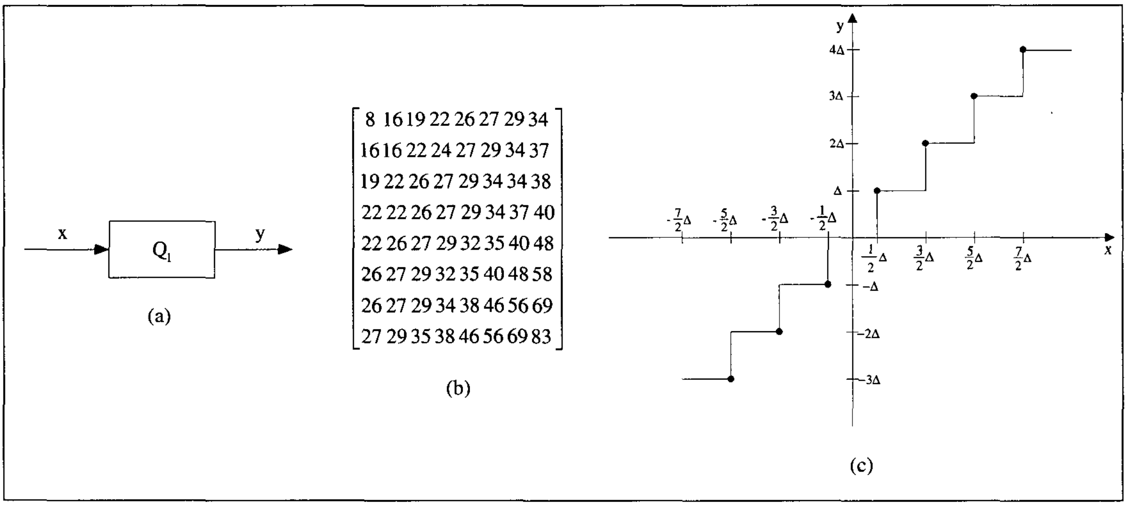
In Fig. 3-(a), the quantizer Q1 maps an original DCT coefficient x into a reconstruction level y, which takes values from a finite set of numbers $\{r_1, \ldots, r_L\}$. The simplified process for this mapping is as follows: define $\{t_m, m = 1, \ldots, L+1\}$ as a set of increasing decision levels with t1 and tL+1 as the minimum and maximum values, respectively, of x. If x lies in the interval $[t_m, t_{m+1})$, then it is mapped to r_m, the m-th reconstruction level.

For a positive value x, the mapping is specified as[3]

$$y = Q_1(x) = \left\lfloor \left\lfloor \frac{x}{\Delta} + \frac{1}{2} \right\rfloor \cdot \Delta \right\rfloor \qquad (2)$$

where $\lfloor a \rfloor$ denotes the largest integer smaller or equal to the given argument a.

If x belongs to the interval $[t_m, t_{m+1})$, it is represented as



(a)

$$\begin{bmatrix} 8 & 16 & 19 & 22 & 26 & 27 & 29 & 34 \\ 16 & 16 & 22 & 24 & 27 & 29 & 34 & 37 \\ 19 & 22 & 26 & 27 & 29 & 34 & 34 & 38 \\ 22 & 22 & 26 & 27 & 29 & 34 & 37 & 40 \\ 22 & 26 & 27 & 29 & 32 & 35 & 40 & 48 \\ 26 & 27 & 29 & 32 & 35 & 40 & 48 & 58 \\ 26 & 27 & 29 & 34 & 38 & 46 & 56 & 69 \\ 27 & 29 & 35 & 38 & 46 & 56 & 69 & 83 \end{bmatrix}$$

(b)

(c)

Fig. 3. Quantizer for MPEG-1 INTRA MB: (a) block diagram (b) quantization weighting matrix (c) quantizer characteristics.

$$x_m = \{x \mid x \in [t_m, t_{m+1})\}. \tag{3}$$

Then the amplitude level to represent x_m is determined as

$$\ell_m = \left\lfloor \frac{x_m}{\Delta} + \frac{1}{2} \right\rfloor. \tag{4}$$

According to Eqs. (2) and (4), x_m is mapped into the reconstruction level r_m

$$r_m = \lfloor \ell_m \cdot \Delta \rfloor. \tag{5}$$

For a negative value x, the mapping can be similarly specified as

$$y = -Q_1(|x|). \tag{6}$$

It can be seen that Q1 is a uniform quantizer without a dead-zone as shown in Fig. 3-(c).

2.2 Quantization Characteristics for INTRA MB of MPEG-4

Fig. 4-(a) shows the block diagram of the quantizer for MPEG-4. Considering the sequential process in transcoding, the input y corresponds to the output of the MPEG-1 quantizer Q1.

A critical difference between Q1 and Q2 occurs from the quantization step size mismatch. Unlike MPEG-1, MPEG-4 simple profile does not employ any weighting matrix for quantization. Thus the quantization step size Δ' for MPEG-4 INTRA MB is defined as[4]

$$\Delta' = 2 \cdot Q_P. \tag{7}$$

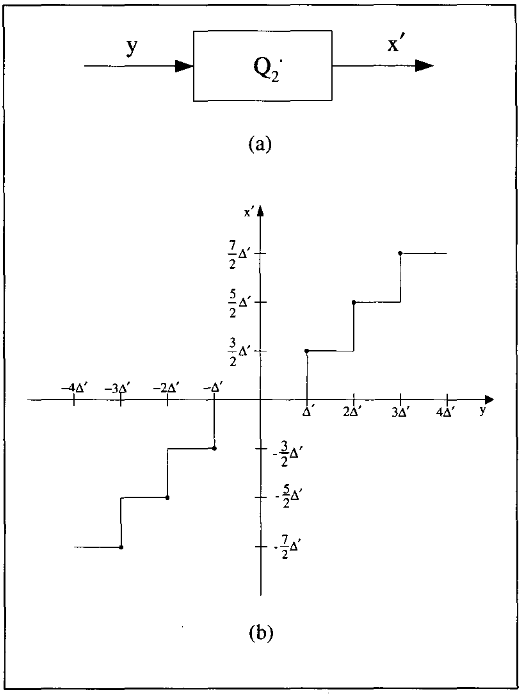
For a positive value y, the mapping between y and x' is specified as[4]

$$x' = Q_2(y) = \begin{cases} \left\lfloor \left\lfloor \frac{y}{\Delta'} \right\rfloor \cdot \Delta' + \frac{\Delta'}{2} \right\rfloor & \text{if } Q_p \text{ is odd,} \\ \left\lfloor \left\lfloor \frac{y}{\Delta'} \right\rfloor \cdot \Delta' + \frac{\Delta'}{2} \right\rfloor - 1 & \text{if } Q_p \text{ is even.} \end{cases} \tag{8}$$

If y belongs to the interval $[t'_n, t'_{n+1})$, we represent it as

$$y_n = \{y \mid y \in [t'_n, t'_{n+1})\}. \tag{9}$$

Then the amplitude level to represent y_n is determined as



(a)

(b)

Fig. 4. Quantizer for MPEG-4 INTRA MB: (a) block diagram of the quantizer (b) quantizer characteristics

$$\ell'_n = \left\lfloor \frac{y_n}{\Delta'} \right\rfloor. \tag{10}$$

According to Eqs. (8) and (10), y_n is mapped into the reconstruction level r'_n

$$r'_n = \begin{cases} \left\lfloor \ell'_n \cdot \Delta' + \frac{1}{2}\Delta' \right\rfloor & \text{if } Q_p \text{ is odd,} \\ \left\lfloor \ell'_n \cdot \Delta' + \frac{1}{2}\Delta' \right\rfloor - 1 & \text{if } Q_p \text{ is even.} \end{cases} \tag{11}$$

The quantizer characteristics is shown in Fig. 4-(b), where we can notice the existence of a dead-zone.

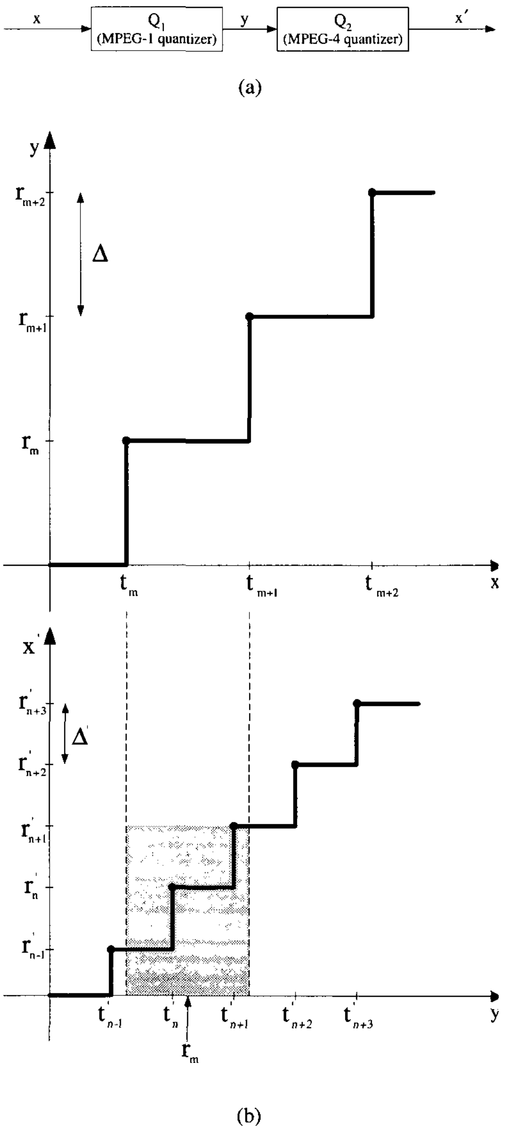
III. Efficient requantization from MPEG-1 to MPEG-4

From the quantization step size mismatch, as evidenced in (1) and (8), Δ is always larger than Δ' if the weighting factor wi is larger than 16. Referring to

Fig. 3-(b), Δ and Δ' are the same only at the frequency index i of 1, 8 and 9 at which the weighting factor is 16. Furthermore the quantizers of MPEG-1 and MPEG-4 are different each other due to the dead-zone mismatch, as evidenced in Fig. 3-(c) and Fig. 4-(b). From these analyses, if requantization is performed independently of the first quantization, the requantization error cannot be avoided.

Fig. 5-(a) shows the sequential quantization process to be considered in the heterogeneous transcoder. Figure 5-(b) shows the quantizer characteristics of the sequential process, where each DCT coefficient x belonging to the interval $[t_m, t_{m+1})$ is mapped into the reconstruction level r_m by the quantizer Q1 and r_m is mapped into the default reconstruction level r'_n after passing through the quantizer Q2. In conclusion, the conventional method simply maps r_m into the default reconstruction level r'_n. However, on closer inspection of Figure 5-(b), there are multiple candidates, such as r'_{n-1}, r'_n and r'_{n+1}, for the final reconstruction level. There is plenty of room for reducing requantization distortion by selecting the more appropriate reconstruction level in the transcoder.

The conventional method simply maps r_m into the default reconstruction level r'_n. In the light of the Lloyd-Max quantizer design, the conventional method is reasonable if the probability distribution of the original DCT coefficient is uniform, where all decision levels as well as reconstruction levels are equally spaced. However, this cannot generally be achieved in practical situations, where the probability distribution of the original DCT coefficient is not uniform. In Lloyd-Max quantizer design, the optimum decision levels lie halfway between the optimum reconstruction levels, which, in turn, lie at the center of mass of the probability density between the decision levels. From the Lloyd-Max quantizer design rule, the PDF of x is introduced to find a near-optimum reconstruction level in the sense of MSE. By using the near-optimum reconstruction level, the mean square error E[(x-x')2] can be reduced when compared with the conventional method.



(a)

(b)

Fig. 5. Sequential quantization process to be considered in the heterogeneous transcoder: (a) block diagram of the sequential quantization process (b) quantizer characteristics of Q1(above) and Q2(below)

For the proposed method, we define, at the following equation, the collection of the subscripts in the decision levels tp' which lie in the half-open interval $[t_m, t_{m+1})$

in Fig. 5-(b) as a set P.

$$\mathbf{P}=\left\{p \mid t'_p \in [t_m, t_{m+1})\right\} \tag{12}$$

The collection of the subscripts to be considered in determining the near-optimum amplitude level is defined as

$$\mathbf{K}=\mathbf{P}\cup\{\min\{\mathbf{P}\}-1\} \tag{13}$$

where the operator $\min\{\mathbf{A}\}$ extracts the minimum element from set A. For the case of Fig. 5-(b), set P is {n, n+1} and set K is {n-1, n, n+1}.

We select an element of K which satisfies the following cost function as the subscript of the near-optimum amplitude level.

$$k=\arg\min_{k\in\mathbf{K}}\left|C_m-r'_k\right| \quad \text{where } C_m=\frac{\int_{t_m}^{t_{m+1}} x\cdot p(x)\,dx}{\int_{t_m}^{t_{m+1}} p(x)\,dx} \tag{14}$$

C_m is the local center of mass in interval $[t_m, t_{m+1})$ and p(x) is the PDF of x. Finally the amplitude level with subscript k, ℓ'_k is selected as the near-optimum amplitude level and is variable length coded in the transcoder.

IV. Method for estimating PDF p(x) in the transcoder

In order to apply the cost function Eq. (16), the transcoder must know p(x). It has been generally believed that the PDF of luminance AC DCT components is Laplacian [1,2]:

$$p(x)=\frac{\lambda}{2}\cdot e^{-\lambda|x|} \tag{15}$$

where a single Laplacian parameter λ determines the characteristics of the distribution of x. Because of the energy compaction property of DCT which packs most of the energy in a few transform coefficients, the distribution for each AC frequency index of an 8×8 block can not be expected to be identical.

Table 1. Laplacian parameters for AC luminance coefficients by Eq. (17).

	0	1	2	3	4	5	6	7
0	—	0.019	0.027	0.040	0.057	0.076	0.107	0.170
1	0.016	0.027	0.035	0.046	0.063	0.088	0.133	0.208
2	0.021	0.029	0.037	0.049	0.066	0.092	0.133	0.220
3	0.025	0.033	0.040	0.050	0.066	0.094	0.148	0.255
4	0.028	0.034	0.044	0.056	0.075	0.107	0.160	0.267
5	0.032	0.037	0.046	0.058	0.078	0.115	0.166	0.293
6	0.036	0.041	0.048	0.062	0.089	0.123	0.173	0.304
7	0.043	0.047	0.056	0.069	0.093	0.131	0.188	0.313

Using Eq. (15), the mean of a random variable |x| is obtained as follows:

$$E(|x|)=\int_{-\infty}^{\infty}|x|\cdot p(x)\,dx=\int_{-\infty}^{\infty}|x|\cdot\frac{\lambda}{2}\cdot e^{-\lambda|x|}\,dx=\frac{1}{\lambda}. \tag{16}$$

As a result of Eq. (16), λ can be simply obtained as

$$\lambda=\frac{1}{E(|x|)}. \tag{17}$$

This computation must be performed for each of the coefficients because they may have different distributions.

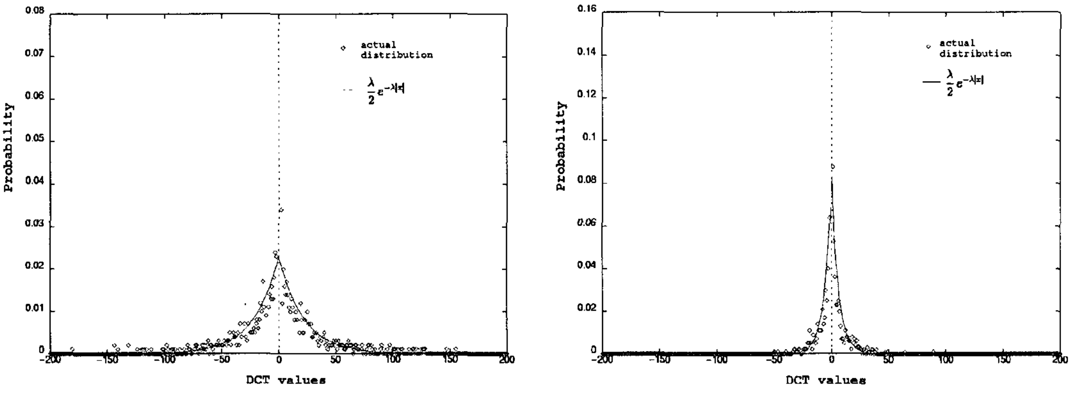
The Laplacian parameters for the luminance coefficients of the first frame of the SIF-formatted flower garden sequence were computed by using Eq. (17). The results are shown in Table 1. As we can expect from the energy compaction property of DCT, the smaller values are generally located at the lower frequency indices.

For verification of the results, Fig. 6-(a) plots the Laplacian distribution with λ=0.046 and the actual distribution of the frequency index of 11 or (1,3) coefficient. Similar plot for the frequency index of 46 or (5,6) coefficient is shown in Fig. 6-(b). From these results, we confirm the validity of the relationship Eq. (17).

Because the information about x in the transcoder is y, which is the reconstructed value of x, we approximate $E(|x|)$ as

$$E(|x|)\cong E(|y|)+E(|y|)_{\left|\frac{\Delta}{2}\right|} \tag{18}$$

where the second term $E(|y|)_{\left|\frac{\Delta}{2}\right|}$ is added to



(a) (b)

Fig. 6. Actual distribution against the Laplacian distribution: (a) Laplacian distribution with λ=0.046 against the actual distribution of (1,3) coefficient, (b) Laplacian distribution with λ=0.166 against the actual distribution of (5,6) coefficient.

compensate for the nonzero x values reconstructed as zero in the transcoder, and defined as

$$E(|y|)_{|\frac{\Delta}{2}|} = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} |y|\cdot p(y)\,dy. \tag{19}$$

The PDF of y needed to compute $E(|y|)_{|\frac{\Delta}{2}|}$ is given by

$$p(y) = \frac{\lambda'}{2}\cdot e^{-\lambda'|y|} \quad \text{where } \lambda' = \frac{1}{E(|y|)}. \tag{20}$$

Using Eq. (20), $E(|y|)_{|\frac{\Delta}{2}|}$ can be obtained as

$$E(|y|)_{|\frac{\Delta}{2}|} = 2\cdot\int_0^{\frac{\Delta}{2}} y\cdot\frac{\lambda'}{2}e^{-\lambda' y}\,dy = \frac{1}{\lambda'} - e^{-\lambda'\Delta/2}\left(\frac{1}{\lambda'}+\frac{\Delta}{2}\right) \tag{21}$$

With Eqs. (18) and (21), it follows from Eq. (17) that λ can be estimated by the following relationship:

$$\begin{aligned}\lambda &= \frac{1}{E(|x|)} \cong \frac{1}{E(|y|)+E(|y|)_{|\frac{\Delta}{2}|}} \\ &= \frac{1}{\frac{1}{\lambda'}+\frac{1}{\lambda'}-e^{-\lambda'\Delta/2}\left(\frac{1}{\lambda'}+\frac{\Delta}{2}\right)} \\ &= \frac{\lambda'}{2-e^{-\lambda'\Delta/2}\left(1+\frac{1}{2}\Delta\lambda'\right)}.\end{aligned} \tag{22}$$

To verify the validity of Eq. (22), a bitstream was generated by encoding the first frame of the SIF-formatted flower garden sequence with the same quantization parameter of 10 within the frame. Then λ values were estimated by using Eq. (22) in the transcoder. The results are shown in Table 2.

Comparison of the parameters in Table 1 and Table 2 shows that the estimated values are very similar to those computed with actual x values. If we employ a rate control where the quantization parameter is changed on a macroblock basis, the quantization parameter for

Table 2. Laplacian parameters for AC luminance coefficients by Eq. (22).

	0	1	2	3	4	5	6	7
0	—	0.018	0.027	0.038	0.055	0.070	0.097	0.169
1	0.016	0.027	0.035	0.043	0.059	0.082	0.121	0.178
2	0.021	0.029	0.036	0.046	0.061	0.083	0.122	0.198
3	0.025	0.032	0.038	0.047	0.060	0.082	0.153	0.224
4	0.028	0.033	0.042	0.051	0.065	0.097	0.177	0.250
5	0.031	0.035	0.044	0.053	0.069	0.105	0.196	0.309
6	0.035	0.039	0.045	0.055	0.076	0.133	0.237	0.415
7	0.042	0.044	0.051	0.061	0.085	0.177	0.223	0.450

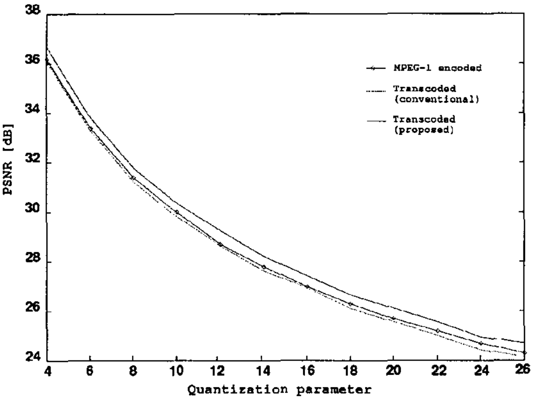


Fig. 7. Average PSNR comparison for flower garden sequence.

the calculation of Δ in Eq. (22) is defined as the average quantization parameter over the frame. Note that the cost function Eq. (14) can be applicable at the cost of only one frame delay for the estimation of λ. Repeated complexity for the calculation of Laplacian parameters for every picture can be avoided comparatively for a long time by using the estimated parameters of the most recently transcoded frame of the same type, since image statistics do not change much over a short period of time.

V. Experimental results

We used the Berkeley MPEG-1 decoder and the MoMuSys MPEG-4 encoder, which are publicly available, to build the heterogeneous transcoder as shown in Fig. 1. The quality of the transcoded sequence measured in PSNR and the amount of the saved bit count were used for evaluation of the proposed method.

The first simulation was carried out on the first ten consecutive frames of the SIF-formatted flower garden sequence. All frames were encoded in the MPEG-1 INTRA mode with the same quantization parameter Qp. Fig. 7 shows the average PSNR values for the transcoded sequence by the proposed method as a function of Qp. As references for the picture quality by the proposed method, the average PSNR values of the original MPEG-1 coded sequence and the transcoded sequence by the conventional method are also shown in the same plot. The PSNR gain of the proposed method over the conventional one ranges from 0.35 to 0.61dB.

Table 3. Average bit counts comparison per frame for flower garden sequence

Quantization parameter	MPEG-1 encoded	Transcoded (conventional)	Transcoded (proposed)	bit saving
4	297509	297456	288873	8583 (2.88%)
6	228122	225486	216834	8652 (3.84%)
8	187950	185324	174358	10966 (5.92%)
10	159041	156820	146432	10388 (6.62%)
12	138606	136491	127892	8599 (6.30%)
14	123148	121285	113698	7587 (6.26%)
16	111746	109453	102143	7310 (6.67%)
18	100669	97342	91624	5718 (5.87%)
20	92440	90138	85690	4448 (4.93%)
22	85294	83487	77847	5640 (6.75%)
24	79598	78400	73643	4757 (6.06%)
26	74153	73924	69320	4604 (6.22%)

Table 3 shows the average bits per frame for the three cases. From this table, we can observe two interesting results. One is that the average bits generated by the proposed method are saved by approximately 2.8~6.7% from the conventional method. We can explain this phenomenon with the statistics of the DCT coefficient and the proposed method to find the near-optimum reconstruction level. As mentioned in the previous section, the probability distribution of the DCT coefficient is Laplacian. With Laplacian distribution, the local center of mass C_m lies on the left of the halfway between tm and tm+1. Because r_m lies halfway between tm and tm+1, C_m is usually smaller than r_m, resulting in the increase of the possibility of selecting smaller reconstruction level as near-optimum. The amplitude level for the smaller reconstruction level guarantees smaller bit generation according to the VLC table of MPEG-4 [4]. Another noticeable result in Table 3 is that the bits after transcoding are always smaller than that of the original MPEG-1 sequence. The reason for this fact is that

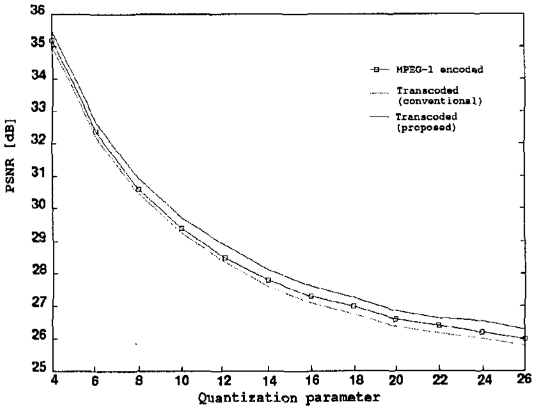


Fig. 8. Average PSNR comparison for table tennis sequence

MPEG-4 simple profile uses the AC-DC prediction algorithm, resulting in enormous bit saving compared to MPEG-1 coding algorithm.

The second simulation was carried out on the first ten consecutive frames of the SIF-formatted table tennis sequence while other simulation conditions are the same as the first simulation. The average PSNR comparison is shown in Fig. 8. The figure shows that the proposed method gives about 0.42~0.55dB improvement in PSNR over the conventional method. Table 4 shows the average bits per frame for the three cases. The bits saving in percentage by the proposed method ranges from 6.4 to 7.4%.

Table 4. Average bit counts comparison per frame for table tennis sequence

Quantization parameter	MPEG-1 encoded	Transcoded (conventional)	Transcoded (proposed)	bit saving
4	172314	168934	157380	11554 (6.83%)
6	124859	120867	115483	8384 (6.76%)
8	98753	94379	88309	6070 (6.43%)
10	80212	77398	72340	5058 (6.53%)
12	67408	64967	60213	4844 (7.45%)
14	58334	55987	51891	4096 (7.31%)
16	52039	49768	46550	3218 (6.46%)
18	45919	43739	40652	3087 (7.05%)
20	41684	39587	36791	2796 (7.06%)
22	38170	36249	33837	2412 (6.65%)
24	35613	33695	31397	2298 (6.82%)
26	33328	31636	29432	2204 (6.97%)

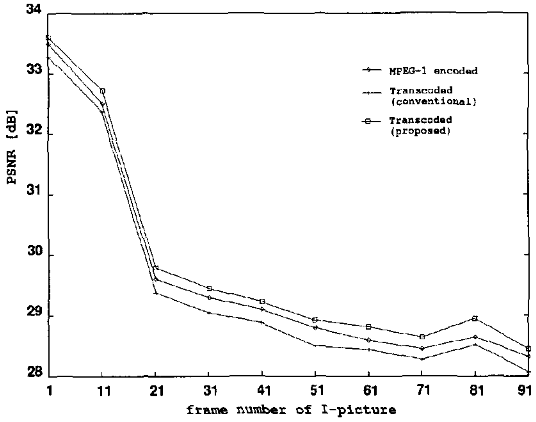


Fig. 9. PSNR comparison for each I-picture

For the third simulation, we generated an original MPEG-1 video sequence by coding the SIF-formatted foot ball sequence at a fixed bit rate of 1.5Mbps. Group of Picture (GOP) size was 10 with structure IPPPPPPPPPI. To investigate the performance of the proposed method, we concentrated only on every 11-th frame of picture type INTRA. PSNR for each 11-th frame by the proposed method is shown in Fig. 9, where a PSNR gain of 0.3~0.5dB over the conventional method can be observed. In this simulation, the average bit count per I-picture was 184373 for the conventional method and 173120 for the proposed method, resulting in up to 6.1% bit saving compared with the conventional approach.

VI. Conclusions

We proposed an efficient requantization method for INTRA MB in heterogeneous transcoding from MPEG-1 to MPEG-4 simple profile where the quantizer mismatch is a critical cause for the quality degradation and excessive bit generation. In order to minimize the quality degradation, the degree of freedom that lies in selecting a near-optimum reconstruction level is exploited. For the efficient requantization method, we have proposed, in the sense of MSE, a novel method to find a near-optimum reconstruction level or equivalently amplitude level. The proposed method requires the statistical property of the original DCT coefficients. Thus we have also presented a simple method for estimating the PDF of the original DCT coefficients based on the Laplacian model. Experimental results show that the proposed method gives 0.3~0.6dB improvement in PSNR over the conventional method, even at the reduced bit-rate about 5~7%.

Although we considered the specific heterogeneous transcoding case from MPEG-1 to MPEG-4, the ideas of this work can also be applicable to other heterogeneous transcoding cases, where the quantizer mismatch is likely to occur.

References

[1] R. C. Reininger and J. D. Gibson, "Distributions of the two-dimensional DCT coefficients for images," IEEE Trans. Commun., vol. com-31, no. 6, pp. 835~839, June 1983.

[2] S. Smoot and L. Rowe, "Study of DCT coefficient distributions,"Human vision and electronic imaging, in proc. of IS&T/SPIE symposium on electronic imaging: science & technology, vol. 2657, San Jose, CA, January 1996.

[3] ISO/IEC 11172-2, Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s, Aug. 1993.

[4] ISO/IEC 14496-2, Coding of audio visual objects: Visual, ISO/IEC JTC1/SC29/WG11, Tokyo, March 1998.

[5] A. K. Jain, Fundamentals of digital image processing, Englewood cliffs, NJ: Prentice Hall, 1989.

[6] N. S. Jayant and P. Noll, Digital coding of waveforms, Englewood cliffs, NJ: Prentice Hall, 1984.

[7] N. Bjork and C. Christopoulos, "Transcoder architecture for video coding," IEEE Trans. Consumer Electronics, vol. 44, no. 1, pp. 88~98, Feb. 1998.

[8] S. Acharya and B. Smith, "Compressed domain transcoding of MPEG," IEEE Int. Conf. Multimedia Computing and Systems, pp. 295~304, Austin, Texas, 1998.

[9] T. Shanableh and M. Ghanbari, "Heterogeneous video transcoding MPEG:1,2 to H.263," Int. Workshop on Packet Video, New York, April 1999.

[10] D. G. Morrison, M. E. Nilsson and M. Ghanbari, "Reduction of the bit-rate of compressed video while in its coded form," Int. Workshop on Packet Video, Portland, Oregon, Sep. 1994.

[11] H. Sun, W. Kwok and J. W. Zdepski, "Architectures for MPEG compressed bitstream scaling," IEEE Trans. Circuits Syst. Video Technol., vol. 6, no. 2, pp. 191-199, April 1996.

[12] G. Keesman, "Transcoding of MPEG bitstream," Signal Process.: Image Commun., Elsevier Science, vol. 8, no.6, pp. 481-500, Sep. 1996.

[13] O. Werner, "Generic quantizer for transcoding of hybrid video," Int. Picture Coding Symposium, pp. 401~406, Berlin, Germany, 1997.

저 자 소 개

Kwang-deok Seo received the B.S. and M.S. degree in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Korea, in 1996 and 1998, respectively. He is currently working toward a Ph.D. degree in electrical engineering at KAIST. He is a student member of IEEE and IEICE. His current research interests include video coding, compressed domain video processing, and video communication over heterogeneous network.

Jae-kyoon Kim received the B.S. degree from National Aviation College, Korea in 1962, the M.S. degree from Seoul National University, Seoul, Korea in 1967, and the Ph.D. degree from the University of Southern California, Los Angeles in 1971, all in electrical engineering. In 1972, he was a Research Associate with NASA/GSFC, Greenbelt, MD. Since 1973, he has been a faculty member of the Korea Advanced Institute of Science and Technology (KAIST), Seoul and Taejon, Korea, where he is a Professor in the Department of Electrical Engineering and Computer Science. He was the president of KICS(Korea Institute of Communications Science) from 1993 to 1994. He is a fellow member of IEE and a member of KICS, IEEK, SMPTE, EURASIP and IEEE. His current research interests are in the areas of video coding and multimedia communication systems.