

Zonal MPEG-2

Cheng-Hsiung Hsieh and Chen-Wei Fu

Department of Computer Science and Information Engineering
Chaoyang University of Technolog
E-mail: chhsieh@cyut.edu.tw

Abstract

In this paper, a simple but effective approach to improve the performance of MPEG-2 is proposed. This study is motivated by (i) the energy compaction property in DCT and (ii) the property of human visual system. Both (i) and (ii) indicate that low-frequency coefficients are more important than others in the DCT transformed blocks. Moreover, the quantized transformed coefficients in high-frequency region are zero or approximate zero in general. Those coefficients contribute nothing or little to PSNR and may waste some bit rate in MPEG-2. Therefore, the zonal coding scheme is applied in the MPEG-2 where some high-frequency coefficients are discarded. The proposed approach is called zonal MPEG-2 which is totally compatible with the traditional MPEG-2. The zonal MPEG-2 is verified by examples and is able to significantly reduce the bit rate with acceptable perceptual quality degradation when compared with the traditional MPEG-2.

1. Introduction

In 1988, the MPEG (Moving Pictures Experts Group) was formed by ISO (International Standard Organization) and IEC (International Electrotechnical Commission) whose objective was to set up video/audio coding standards. One of standards pervades in the coding community is the MPEG-2 which was announced in 1994. The MPEG-2 standard has been adopted in HDTV and DVD and has great impact on consumer electronics. Therefore, many schemes to improve the performance of MPEG-2 have been reported recently. In [1-2], the performance of MPEG-2 is improved through rate control mechanism. In [3], SNR scalability and stream morphing are achieved by rate control. In [4], a hybrid error concealment scheme is proposed for MPEG-2. In this paper, an approach for MPEG-2 to trade some PSNR with significant bit rate reduction is proposed. This paper is organized as follows: Section 2 gives a brief review of the MPEG-2. Then the proposed approach based on the zonal coding scheme is described in Section 3. Section 4 justifies the proposed approach through examples where discussions are given as well. Finally, conclusion is made in Section 5.

2. Review of the MPEG-2

In this section, the coding process of MPEG-2 is briefly reviewed. For details, one may consult [5-7]. The structure of MPEG-2 is shown in Figure 1. In the MPEG-2, the spatial redundancy is reduced by DCT (Discrete Cosine Transform) while the temporal redundancy is eliminated by motion estimation and motion compensation.

In the MPEG-2, the input image sequence is first put into the unit of images reorder. And GOP (Group of Pictures) is formed by an I-frame (Intra Frame), several P-frame (Predictive Frame) and B-frame (Bidirectionally-Predictive Frame). In the coding process, the image is divided into macroblocks which is the

basic unit for motion estimation and motion compensation. In the case of 4:2:0 sampling, a macroblock consists of four luminance blocks (Y-block) and two chroma blocks (U-block and V-block). The size for Y-, U-, and V-block is 8×8 which is the basic unit for the transform coding in the MPEG-2. Two types of block are involved in the coding process: intra-block and inter-block. The relationship of I-, P-, B-frame and intra-block, inter-block is given in Figure 2 and described in the following.

2.1 Frames and blocks

(A) I-frame and intra block

An intra-block uses only its own data to reduce the spatial redundancy. All blocks in I-frame are intra-block. The intra-block requires no reference frame. Thus it can be coded and decoded independently.

(B) P-, B-frame and intra-block, inter-block

In the coding process, P- and B-frame need the preceding or/and the behind frame as a reference frame. For every macroblock, the unit of motion estimation tries to match similar macroblock in the reference frame. When a match is found, the blocks in the macroblock are considered as inter-block and the unit of motion compensation is performed. Otherwise, the blocks in the macroblock are treated as intra-blocks.

2.2 Block coding

Assume that the number of images in the input image sequence, as shown in Figure 2, is K and each image O_i is divided into N_b image blocks $\{b_{ij}\}$ for $1 \leq i \leq K, 1 \leq j \leq N_b\}$ of size 8×8 . Then intra-block and inter-block coding are described in the following.

(A) intra-block coding

- Step 1. Perform DCT on b_{ij} , $B_{ij} = DCT(b_{ij})$, where $DCT(\cdot)$ denotes as discrete cosine transform [8].
- Step 2. Quantize B_{ij} as $\bar{B}_{ij} = B_{ij} ./ Q$, where Q is a quantization matrix and $./$ is the element to element division.
- Step 3. Reorder elements of $\bar{B}_{ij}$ by a scan order. Let $SO[k]$ be the index of the scan order. Then the reordered elements of $\bar{B}_{ij}$ are denoted as $\bar{B}_{ij,SO}[k] = \bar{B}_{ij}[SO[k]]$.
- Step 4. Encode $\bar{B}_{ij,SO}$ by RLC (Run Length Coding) and then VLC (Variable Length Coding).

(B) inter-block coding

- Step 1. Perform the motion estimation. If there is no match for b_{ij} and reference block b_{ref} , code b_{ij} by the intra-block coding. Otherwise, calculate the motion vector as $MV = Vector(b_{ij} - b_{ref})$, where $Vector(\cdot)$ is a function to compute vector. Then put MV to VLC unit.
- Step 2. Find the difference between b_{ij} and b_{ref} as $\tilde{b}_{ij} = b_{ij} - b_{ref}$ and perform the unit of motion compensation.
- Step 3. The rest of steps are identical to the intra-block coding except $b_{ij} = \tilde{b}_{ij}$.

In the coding of MPEG-2, the unit of rate control trims the bitstream by adjusting the quantization step. By this, a bit rate control in a transmission is achieved.

3. The Proposed Zonal MPEG-2

In this section, the motivation for the proposed zonal MPEG-2 is described in Section 3.1. Then, the implementation steps for the zonal MPEG-2 are given in Section 3.2. Finally, an example is provided to demonstrate how the zonal MPEG-2 works.

3.1 Motivation

The proposed zonal MPEG-2 is motivated by the following two properties: (i) the energy compaction property in DCT and (ii) the sensitivity property in human visual system [9]. For the intra-block coding, the energy distribution of DCT-transformed block descends from low-frequency to high-frequency coefficients in general. And for the inter-block coding, a similar distribution as the intra-block coding results because of motion compensation [5]. On the other hand, according to the human visual system, human eyes are more sensitive to low-frequency than to high-frequency. Therefore, low-frequency coefficients are finely quantized and those in high-frequency region are quantized coarsely. Consequently, after quantization high-frequency coefficients become zero or small. There is a high possibility that small quantized coefficients may locate between zero coefficients. These non-zero coefficients generally contribute little PSNR but waste many bits in the coding process. By the ideas just described, we present a simple but effective approach to improve the performance of MPEG-2 where some quantized coefficients in the high-frequency are discarded. Since the concept in the proposed approach is identical to the zonal coding scheme [10-11]. It is thus called the zonal MPEG-2.

3.2 Zonal MPEG-2

In the proposed approach, a parameter M is used to determine the number of coefficients to be coded in the MPEG-2. As mentioned previously, this is the scheme in the zonal coding. Thus, the proposed approach is called the zonal MPEG-2. The coding process for zonal MPEG-2 is identical to the traditional MPEG-2 except $\bar{B}_{ij,SO}[k]$ in Step 3 in the intra-block coding is modified as follows:

Step 3'. Given M , modify $\bar{B}_{ij,SO}[k]$ as

$$\bar{B}_{ij,SO}[k] = \begin{cases} \bar{B}_{ij}[SO[k]], & k \leq M \\ 0, & k > M \end{cases}$$

where $1 \leq M \leq 64$.

It should be pointed out that the zonal MPEG-2 is identical to the traditional MPEG-2 when $M = 64$. Therefore, it is totally compatible with traditional MPEG-2. Moreover, it is found that the inter-block coding affects PSNR significantly since several frames are involved in the inter-block coding. Consequently, only intra-blocks are involved in the proposed zonal MPEG-2.

3.3 An example

Here, an example is provided to demonstrate how the zonal MPEG-2 works. Suppose an 8×8 quantized transformed block is found as shown in Figure 3(a) in the coding process. After RLC, the result is given in Figure 3(b). Then the result of RLC is put into VLC whose output is shown in Figure 3(c). If the first 32 coefficients, i.e., $M = 32$, is selected and coded. The coding result is given in Figure 3(d). To compare Figure 3(c) with Figure 3(d), 22 bits are saved in the zonal MPEG-2 with two discarded coefficients.

4. Simulation Results and Discussions

In this section, the proposed zonal MPEG-2 is verified and compared with the traditional MPEG-2. Video files Foreman and Carphone are used as examples. The information about these files is summarized in Table 1. In the simulation, the software MPEG-2 Test Model 5 (TM5) [12] is employed where the mail profile is considered. Moreover, for fair comparison the rate control is fixed, i.e., the parameter $mquant$ in TM5 set to a constant, in all simulations on purpose. By the traditional MPEG-2, an average PSNR, $PSNR_{avg}$, and file size, $FSize$, for Foreman video are 45.9 dB and 6,552,816 bytes, respectively while $PSNR_{avg} = 44.4$ dB and $FSize = 457,374$ bytes for Carphone video are obtained.

4.1 Effect of M

Two cases are considered to investigate the effect of M on the coding performance of zonal MPEG-2. In the first case, all intra-blocks, except the Y-block in I-frame, are coded by the zonal MPEG-2 where same M is used. The simulation results are given in Table 2 where different values of M are used, and the corresponding differences, $\Delta PSNR_{avg}$ and $\Delta FSize$, for the zonal and traditional MPEG-2 are given.

The second case is same as the first case except the Y-block in the I-frame is put into account. Besides, only Foreman example is considered here. In the simulation, M is set to 8, 16 and M_{I-Y} to 8, 16, 24, 32, 40, 48, 56, respectively. Notation M_{I-Y} is the number of coefficients coded in the Y-block of I-frame. The simulation results for different M and M_{I-Y} are recorded in Table 3 where $\Delta PSNR_{avg,8}$, $\Delta FSize_8$, and $\Delta PSNR_{avg,16}$, $\Delta FSize_{16}$ stand for the differences of the first case with $M = 8$, $M = 16$ from the second case with different M_{I-Y} .

4.2 Discussions

The simulation results shown in Tables 2 and 3 indicate that parameter M is able to reduce $FSize$ with some $PSNR_{avg}$ loss. For example, in the first case, when compared with the traditional MPEG-2, $\Delta PSNR_{avg} = 3.6$ dB and $\Delta FSize = 219,974$ bytes (about 3.4% reduction) are achieved for the Foreman video, when $M = 8$. This is also true in the second case. In the case of $M = 8$ and $M_{I-Y} = 32$, $\Delta PSNR_{avg} = 5$ dB and $\Delta FSize = 339,809$ bytes (about 5.2% reduction) are obtained when compared with the traditional MPEG-2.

Though $\Delta PSNR_{avg}$ are as high as 3.6 dB and 5 dB in the two cases just mentioned, however the perceptual quality of the reconstructed frames degrades very little. Figure 4 shows the first frames of Foreman video obtained from (i) original image, (ii) traditional MPEG-2, (iii) zonal MPEG-2 with $M = 8$, and (iv) zonal MPEG-2 with $M = 8$ and $M_{I-Y} = 32$. From Figure 4, it can be noticed that the reconstructed images of the zonal MPEG-2 has little difference from that for the traditional MPEG-2.

As shown in Table 3, M_{I-Y} has significant effect on $PSNR_{avg}$ and $FSize$. Therefore, as a rule of thumb the combination of $M = 8$ or $M = 16$ with $M_{I-Y} = 32$ is suggested for the proposed zonal MPEG-2. In this case, the perceptual quality degradation is little and the file size reduction is significant.

5. Conclusion

The paper presents a simple but effective approach to improve the performance of traditional MPEG-2. The proposed approach is called zonal MPEG-2 since the zonal coding scheme is implanted into the traditional MPEG-2. Two observations lead to the zonal coding scheme. The first one is the property of energy compaction in DCT. The second is the property in the human visual system. Both properties indicate that low-frequency coefficients are more important than those in high-frequency. Consequently, the perceptual quality should not degrade too much when some high-frequency coefficients are discarded. By truncating some high-frequency coefficients, it is able to reduce the file size of compressed video with some PSNR loss. This is the fundamental idea in the proposed zonal MPEG2. To verify the zonal MPEG-2, two video files are given as examples. The simulation results indicate that the zonal MPEG-2 is able to trade little perceptual quality with significant file size reduction when compared with the traditional MPEG-2.

References

- [1] B. C. Song and K.-W. Chun, "A Virtual Frame Rate Control Algorithm for Efficient MPEG-2 Video Encoding", *IEEE Transactions on Consumer Electronics*, Vol. 49, No. 2, pp. 460-465, May 2003.
- [2] W. Limin, "Rate Control for MPEG Video Coding," *Signal Processing: Image Communication*, Vol. 15, No. 6, pp. 493-511, March 2000.
- [3] M. James, F. Michael and A. John, "Rate Control for MPEG-2 SNR Scalability and Stream Morphing Using Codeword Estimation and Overhead Modeling," *Signal Processing: Image Communication*, Vol. 18, No. 9, pp. 837-853, Oct. 2003.
- [4] Li-Wei Kang and Jin-Jang Leou, "A Hybrid Error Concealment Scheme for MPEG-2 Video Transmission Based on Best Neighborhood Matching Algorithm," *Journal of Visual Communication and Image Representation*, Vol. 16, No. 3, pp. 288-310, June 2005.
- [5] Iain E.G. Richardson, *Video Codec Design*, John Wiley & Sons Ltd, 2003.
- [6] P.N. Tudor, "MPEG-2 Video Compression," *Electronics & Communication Engineering Journal*, Vol. 7, No. 6, pp. 257-264, Dec. 1995.
- [7] ISO-IEC/JTC1/SC29/WG11, "Generic Coding of Moving Pictures and Associated Audio Information: Video," ISO-IEC 13818-2, Nov. 1994.
- [8] K.R. Rao and P. YIP, *Discrete Cosine Transform: Algorithms, Advantages, Applications*, Academic Press, 1990.
- [9] B. Chitprasert and K.R. Rao, "Human Visual Weighted Progressive Image Transmission" *IEEE Trans. on Communications*, Vol. 38, No. 7, pp. 1040-1044, July 1990.
- [10] A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, 1989.
- [11] R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, 2nd Edition, Prentice Hall, 2002.
- [12] ISO/MPEG II, Test Model 5, April 1993. (<http://www.mpeg.org/MPEG/video.html>)

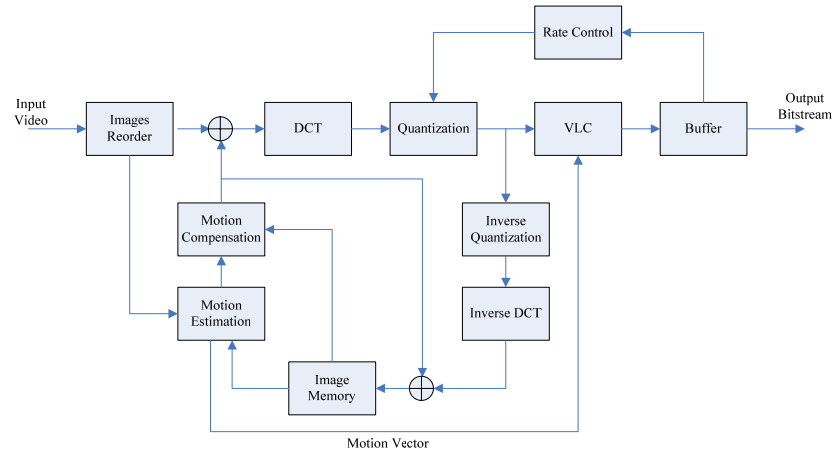


Figure 1. The structure of MPEG-2

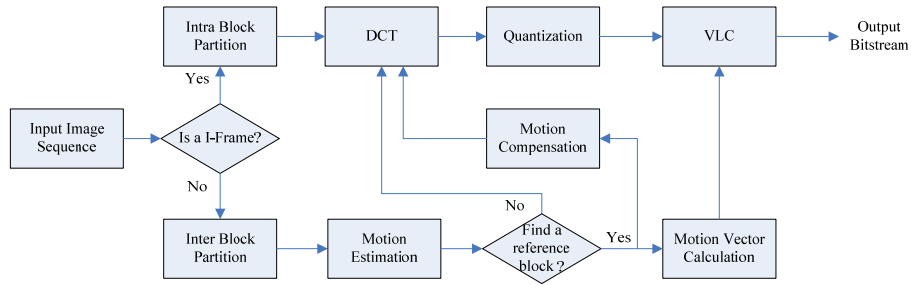


Figure 2. The coding process for intra-block and inter-block

[-26 -3 1 -3 -2 -6 2 -4 1 -4 1 1 5 0 2 0 0 -1 2 0 0
0 0 0 -1 -1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 -1 EOB]

(a)

<0,-26><0,-3><0,1><0,-3><0,-2><0,-6><0,2><0,-4><0,1><0,-4>
<0,1><0,1><0,5><1,2><2,-1><0,2><5,-1><0,-1><7,1><6,-1><EOB>

(b)

[0100 001 0100 0101 100001 0110 100011 001 100011 001 001 100101
11100110 110110 0110 11110100 000 111110011 11110110 1010]

(c)

[0100 001 0100 0101 100001 0110 100011 001 100011 001
001 100101 11100110 110110 0110 11110100 000 1010]

(d)

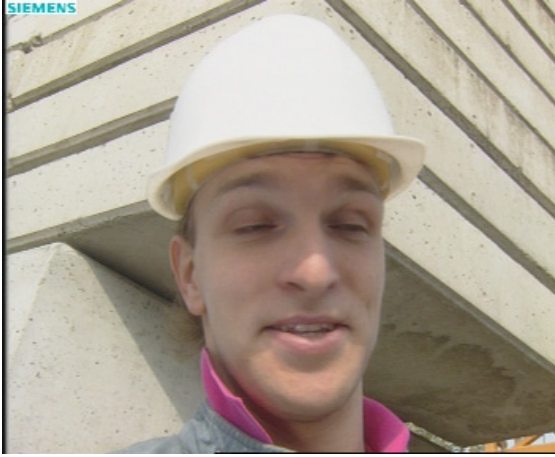
Figure 3. An example of the zonal MPEG-2

(a) Quantized coefficients after scan reordering

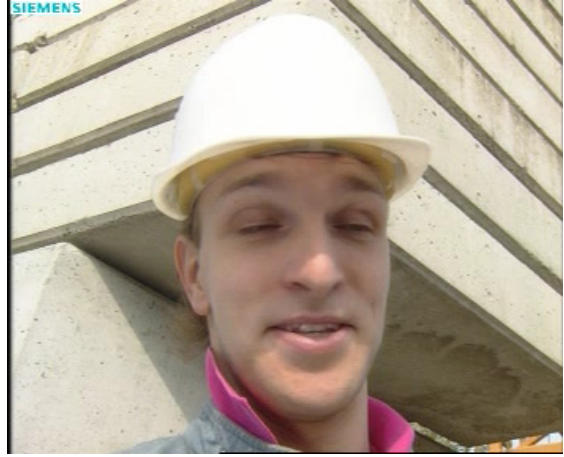
(b) Result after RLC

(c) Result after VLC

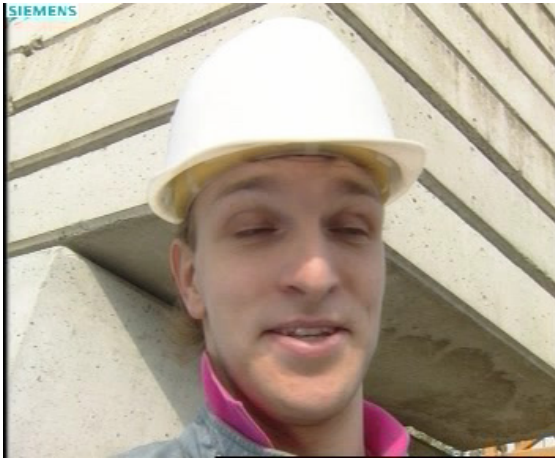
(d) Result after VLC with truncation ($M = 32$)



(a)



(b)



(c)



(d)

Figure 4. The first frames in Foreman video
 (a) original
 (b) reconstructed from the traditional MPEG-2
 (c) reconstructed from the zonal MPEG-2 ($M = 8$)
 (d) reconstructed from the zonal MPEG-2 ($M = 8$ and $M_{I-Y} = 32$)

Table 1. Information for video files used as examples

Filename	Format	Resolution	Number of Frames	File Size (Byte)
Foreman	SIF	352×288	300	45,619,200
Carphone	QCIF	176×144	96	3,649,536

Table 2. Effect of M on $\Delta PSNR_{avg}$ and $\Delta FSize$

M	Foreman		Carphone	
	(45.9 dB) $\Delta PSNR_{avg}$	(6,552,816 Bytes) $\Delta FSize$	(44.4 dB) $\Delta PSNR_{avg}$	(457,374 Bytes) $\Delta FSize$
8	3.6	219,974	4.0	18,126
16	1.6	135,352	2.2	12,191
24	0.8	72,586	0.9	7,630
32	0.5	48,416	0.6	5,016
40	0.2	19,530	0.2	5,016
48	0.1	6,792	0.08	1,015
56	0.007	1,415	0.02	253

Table 3. Effect of M and M_{I-Y} on $\Delta PSNR_{avg}$ and $\Delta FSize$

M_{I-Y}	Foreman			
	$M = 8$		$M = 16$	
	(42.3 dB) $\Delta PSNR_{avg,8}$	(6,332,842 Bytes) $\Delta FSize_8$	(44.3 dB) $\Delta PSNR_{avg,16}$	(6,417,464 Bytes) $\Delta FSize_{16}$
8	3.7	370,378	3.9	370,377
16	2.7	263,135	2.9	263,125
24	1.8	169,148	1.9	169,144
32	1.4	119,835	1.5	119,819
40	0.7	57,498	0.7	57,498
48	0.4	23,602	0.4	23,605
56	0.1	5,879	0.1	5,876