HYBRID-MODE EMBEDDED COMPRESSION FOR H.264/AVC VIDEO CODING SYSTEM

Tung-Chien Chen, Yu-Han Chen, Ke-Chung Wu, Liang-Gee Chen

DSP/IC Design Lab, Graduate Institute of Electronics Engineering and
Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan
{djchen,doliamo}@video.ee.ntu.edu.tw, ddddog@gmail.com, lgchen@video.ee.ntu.edu.tw

ABSTRACT

Applications of high resolution videos have great potential for H.264/AVC. However, due to the multi-frame motion estimation and large search range (SR) requirement, the ultra high system bandwidth becomes the challenge for the plateform based video codec. In this paper, a hybrid-mode embedded compression (EC) is proposed. Two different strategies are respectively used to compress the reconstructed macroblocks (MBs) of intra- and inter-mode. Up-to 9.2 times of compression ratio (CR) can be achieved even with lossless-compression constraint. Besides, with resource sharing, this system can be integrated into H.264/AVC codec with almost no area overhead. According to the simulation results, the system bandwidth can be reduced by 66.2% and 75.3% in average for high and median quality situation.

1. INTRODUCTION

H.264/AVC [1] is the new video coding standard developed by ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG). It can save 25%-45% and 50%-70% of bitrates compared with MPEG-4 advanced simple profile and MPEG-2, respectively. The coding gain mainly comes from new prediction tools, and enormous computation and ultra high memory bandwidth are the penalties. According to instruction profiling, 2.76 tera-operations/s of computational loading and 4.25 tera-bytes/s of memory access are required for real-time encoding SDTV (YUV420, 720x480, 30fps) videos (JM8.5 [2], baseline options, full search, four reference frames, SR [-32, +31]). For platform-based VLSI systems in which the high computation requirement can be easily solved by increasing the parallelism of processing elements, the real challenge is the unacceptable bus bandwidth requirement with limited system resources.

The bandwidth mainly comes from the access of reference data during motion estimation (ME). One common solution is to use the EC for the frame buffer access [3]. The EC engine compresses the MBs of the reconstructed frame and transmits such bitstream to frame buffer. When the video codec system performs ME, the EC engine fetches and decodes the compressed data from system bus. Depending on the target applications, there are various EC algorithms with two categories. For lossy ECs [4][5], they have better performance in bandwidth reduction, but quality degradation occurs due to error propagation. For lossless ones [6] that guarantee the highest video quality, their CRs are limited to two, just like lossless image coding [7]. In this paper, a novel EC system is proposed. Unlike previous ECs that equally process every MB, two different strategies are used for inter- and intra-mode reconstructed MBs, respectively. By this way, our EC can losslessly compress the reconstruction MB with higher CR.

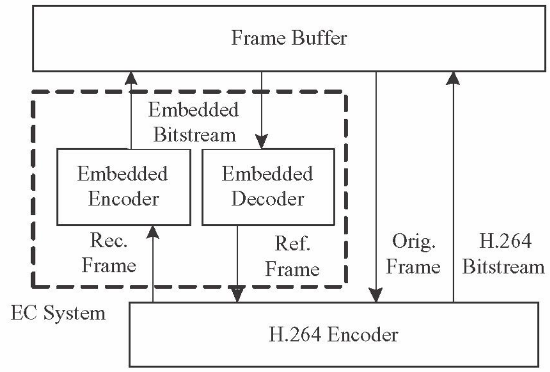

Fig. 1. The EC system for H.264 video encoder.

The rest of this paper is organized as follows. In Section 2, the bandwidth problem is described followed by the analysis of the required EC system. Then, the hybrid-mode EC system is proposed in Section 3. The corresponding hardware architecture as well as performance evaluation are presented in Section 4 and Section 5. Finally, Section 6 gives a conclusion.

2. PROBLEM STATEMENT

In ME, in order to find the best matched candidate, a search window (SW) within one reference frame has to be searched. A huge amount of reference data must be loaded from frame buffer to ME core, and the traffic is very heavy. Because pixels in neighboring candidate blocks are considerably overlapped, and so are the SWs of neighboring current MBs, the bandwidth of system bus can be greatly reduced by designing local buffers to store reusable data [8][9]. By means of local memory access, the external memory bandwidth can be reduced.

For high resolution video application in H.264, however, the bandwidth is still too large even with the above techniques. The larger frame size means that the larger SR is required to achieve good ME performance. The system bandwidth would exponentially increase. Besides, H.264 support the feature of multiple reference frames [10]. The required reference data are proportional to the number of reference frames, which also heavily increases the system bandwidth. Table 1 and Table 2 show the necessary system bandwidth of ME for different video formats and reference frame numbers, respectively. The analysis is based on level-C data reuse strategy [9] that is adopted the most frequently nowadays. The system bandwidth increases exponentially with the frame size and

0-7803-9266-3/05/$20.00 ©2005 IEEE.

Table 1. The necessary system bandwidth of ME for different frame size with one reference frame.

Format	SR	Bandwidth
CIF(352×288, 30fps)	±16	9.13MB/s
D1(720×480, 30fps)	±64	93.3MB/s
HDTV(1280×720, 30fps)	±128	470MB/s
HDTV(1920×1080)	±256	2.05GB/s

Table 2. The necessary system bandwidth of ME for different multiple reference frames (HDTV, 1280×720, 30fps with ±128 SR).

# of ref. frame	system bandwidth
2	940MB/s
3	1.41GB/s
4	1.88GB/s

linearly with reference frame number. In the HDTV case with 4 reference frames, the bandwidth requirement is not achievable for plateform based system. To do further reduction, EC techniques should be used. Figure 1 shows the EC system for H.264 video encoder. The EC compresses the reconstructed frame data into embedded bitstream before sending out , and then the compressed data will be read and decoded while the SW is loaded. Since only the compressed bitstream is transmitted on system bus, the bandwidth can thus be reduced. Several issues are considered for our EC system in H.264 large frame size application:

1. Lossless compression: Most of EC for previous standards are lossy. The lossy compression has better CR at an expense of video quality. Some problems such as data mismatch between encoding and decoding may occur, which induce the error propagation. Therefore, due to the demand of high video quality in the HDTV application, the compressed reference frame must be able to perfectly reconstructed.
2. High CR: The system bandwidth reduction ratio is equal to the CR of EC. The encoding efficiency must be high enough to solve the large system bandwidth problem. In general, the CR limitation of lossless image coding is two. However, EC compresses the reconstructed frame data which have been processed through DPCM loop by original encoder. With the information from original encoder , our EC can achieve higher CR.
3. Resource sharing: There are many encoding schemes in H.264 encoder with high compression efficiency. Some of them can be chosen as parts of our EC. We can share existed functional blocks in H.264 codec to reduce hardware cost for EC system.

3. PROPOSED ALGORITHM

In this section, a new hybrid-mode EC system for H.264 is proposed to meet the required issues mentioned in the previous section. In H.264, a MB is either intra mode or inter mode. For

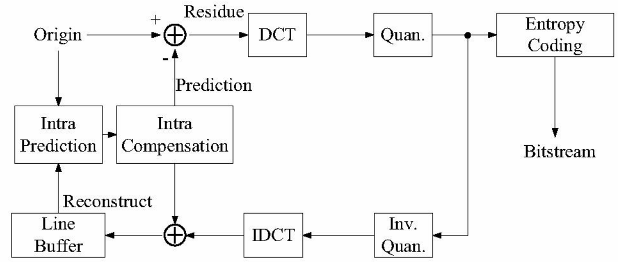

Fig. 2. The H.264 Intra-frame coding.

each mode, different strategies will be used to compress the reconstructed data of H.264.

3.1. Intra mode MB

The H.264 encoder converts the original frames to bitstream, and this bitstream can be converted to reconstructed frames by H.264 decoder. Therefore, though the H.264 bitstream is lossily compressed data for original frames, it is losslessly compressed ones for reconstruction frames. Based on this idea, EC system may directly use the H.264 bitstream as the embedded bitstream, which will be transmitted to and stored in the external memory. During ME for the next original frame, the reference frame, the previous reconstruction frame, can thus be perfectly recovered from this bitstream. However, this scheme sometimes does contradict the demands of EC. If a MB is inter-mode encoded in H.264, the temporal information is required when we reconstruct it. In this way, more data are involved during loading search window, and the system bandwidth cannot be reduced.

In order to properly solve the bandwidth problem, only spatially local data can be involved in EC system, just like image coding. For an intra-mode coded MB, both the encoding and decoding processes access only local information in intra-prediction. Therefore, the H.264 bitstream of intra-mode coded MB can directly use as embedded bitstream without any bandwidth overhead. Please note that, the quantization step in intra-mode encoding reduces the energy and results in better CR. Since our lossless EC is embedded with this quantization step, the CR, or bandwidth reduction ratio, is much higher than that of other EC systems using lossless compression image coding standards. Consider a special case : the reference frame of the first P frame is I frame where all MBs are intra-mode coded. The memory reduction ratio is equal to the H.264 CR, which is much higher than the limitation of the lossless image coding. For general case, the overall CR of our EC system will depend on the percentage of intra-mode coded MBs.

3.2. Inter mode MB

For the inter-mode coded MB, the extra encoding schemes are proposed for our EC system. The H.264 intra-frame coding is chosen as the prototype of the embedded encoder under the considerations of good compression capability, low memory usage, and high resource-sharing possibility. Figure 2 shows the original block diagram of H.264 intra-frame coding. For lossless consideration, the quantization and inverse quantization are removed firstly. Besides, in lossless EC system, since the reconstructed frame must be the same with the original one, the feedback path of DPCM loop can be omitted. After the above modifications, only the transformation and entropy coding are preformed after intra prediction and intra compensation.

We analysis three different strategies of transform inherited from H.264 encoder. The selection depends on both the compression performances and hardware considerations.

Scheme 1: The transform is simply the original integer discrete cosine transform (DCT) in H.264 DPCM loop.

Scheme 2: Hadamard transform, the other transformation in H.264 involved in both the DPCM loop and the encoder issues, is used.

Scheme 3: No transformation. The residue is directly bypassed and entropy coded.

In scheme 1 and 2, the residues are transformed from spatial domain to frequency domain. We must find an inverse transformation that the residues can be losslessly recovered. It means, for any matrix A, if T is the transform matrix, an inverse transform matrix R will satisfy

$$R(TAT^T)R^T = A \tag{1}$$

We separate R into two matrixes, and the equation becomes

$$(B(D(TAT^T)D^T)B^T) = A \tag{2}$$

In scheme 1, T_{scheme_1} is DCT matrix and D_{scheme_1} is IDCT matrix. In order to satisfy (1), the additional matrix of B_{scheme_1} must be implemented as

$$B_{scheme_1} = \begin{pmatrix} 1/4 & 0 & 0 & 0 \\ 0 & 1/5 & 0 & 0 \\ 0 & 0 & 1/5 & 0 \\ 0 & 0 & 0 & 1/4 \end{pmatrix}$$

In scheme 2, T_{scheme_2} is Hadamard transform matrix and D_{scheme_2} is inverse Hadamard transform matrix, that is the same as Hadamard transform matrix. Similarly, in order to satisfy (1), the matrix of B_{scheme_2} is

$$B_{scheme_2} = \begin{pmatrix} 1/4 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1/4 & 0 \\ 0 & 0 & 0 & 1/4 \end{pmatrix}$$

The matrix of B_{scheme_1} is not suitable for hardware implementation due to the requirement of the divisors. On the contrary, the matrix of B_{scheme_2} only involves the shifter, and no hardware overheads are required. Therefore, the scheme 2 is a better choice.

Since both the scheme 2 and scheme 3 have no hardware overhead, they are compared according to the compression performance. Table 3 shows the simulation results. Several video sequences are simulated with quantization parameter (QP) set to 15. According to the experiment, the CR is better without transformation. That is because there is no quantization process in the EC, and transformation can not successfully gather the energy up. Then, scheme 3 is the best choice among all schemes. It is worth to note that the CR of EC system increases with the larger frame size and the higher QP value of the original encoder.

There is another hardware issue included in the proposed algorithm. The correlation of the best intra prediction mode between the original MB of original encoder and the reconstructed MB of embedded encoder is quite high. Therefore, after the intra prediction in the H.264 encoder, the best intra mode information can be reused for our embedded encoder even if the final mode of current coded MB is not intra mode. The intra prediction occupies most of the computation complexity in the embedded encoding. With such modification, the embedded encoder can be implemented with less hardware resource.

Table 3. The simulation results of CR in scheme 2&3 (frame size: (1)352x288 (2)720x288 (3)720x480 (4)750x576 (5)1920x720).

Image name (QP=15)	CR	
	scheme 2	scheme 3
Foreman(1)	1.515	**1.567**
Stefan(1)	1.197	**1.230**
Boat(2)	1.460	**1.492**
Bamboo(3)	2.415	**2.439**
Wendy(4)	1.574	**1.618**
Bigships(5)	2.309	**2.347**

3.3. Hybrid-mode EC system

There are two modes in the proposed EC system : inter mode and intra mode. The mode selection depends on the mode decision of the H.264 encoder. The inter-predicted MB enters the inter mode, while the intra-predicted MB enters the intra mode. For inter mode, we convert the reconstructed MB to embedded bitstream by the proposed EC algorithm. For intra mode, the embedded encoder is idle, and the corresponding H.264 bitstream can be reused as embedded bitstream.

4. ARCHITECTURE

4.1. Embedded encoder

Figure 3 shows the architecture of embedded encoder. If the current MB is inter-mode coded MB, the embedded encoder will compress the reconstructed MB into embedded bitstream. Because the required local information have been prepared during the intra prediction process of the original encoder, and so does the best intra mode, only the intra compensation and entropy coding is involved in our embedded encoder. Reconstructed pixels are subtracted by predictors generated by intra compensation, and residues are then coded by entropy coding module. Please note that these two modules can be shared with original encoder with adequate schedule. Both the hardware cost and computation of the embedded encoder are small. If the current coded MB is intra mode MB, the original bitstream of this MB will be directly used as the embedded bitstream. The embedded encoder is idled in this situation.

4.2. Embedded decoder

During loading SW data, the corresponding embedded bitstream is read from system memory and then decoded by embedded decoder. The decoding procedure is the reverse process of the encoding one. Figure 4 illustrate the architecture of the embedded decoder. The header in embedded bitstream is decoded firstly, and the mode of the loaded MB can be decided. If the MB is intra mode MB, the residues are processed through IQ/IDCT and then added by predictors generated by intra compensation. Otherwise, the bypassing path in the bottom is chosen. Similarly, the hardware

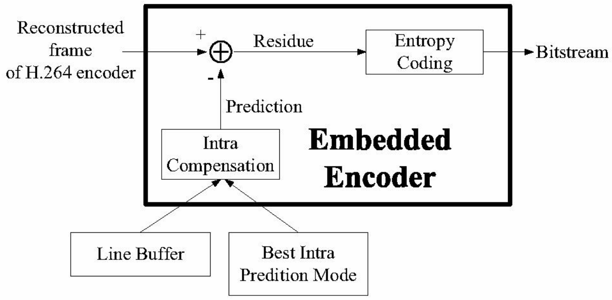

Fig. 3. Proposed architecture of embedded encoder.

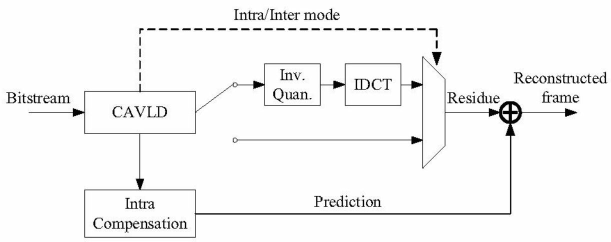

Fig. 4. Proposed architecture of embedded decoder.

source sharing can also be achieved if the schedule of the coding process is carefully designed.

5. SIMULATION RESULTS

The performance of the proposed hybrid-mode EC is shown in Table 4. The reference software, JM8.5 [2], is modified, and four sequences with SDTV or HDTV formats that match the target applications are selected for the simulation. Two quantization parameters standing for high and median quality situations are used. The CR can be obtained by the following equation:

$$\frac{1}{CR_{hybrid}} = \frac{\alpha}{CR_{intra}} + \frac{1-\alpha}{CR_{inter}}$$

where CR_{intra}: CR for intra-mode MB.
CR_{inter}: CR for inter-mode MB
α : percentage of intra predicted MB

The CR of the intra mode MB's is much better than that of inter mode ones. The overall CR will be largely influenced by the percentage of the intra mode MB's. The overall CR, or the bandwidth reduction ratio, can be as high as 9.259. In worst case with almost no intra MB, the CR would approach to two, the upper-bound of lossless image coding. With our approach, the system bandwidth can be reduced by 66.2% and 75.3% in average for high and median quality situations, respectively. Please note that we didn't do any modification on entropy coding for resource sharing. The performance of the proposed EC can be better if appropriate revision is made according to the statistics.

6. CONCLUSION

In this paper, a hybrid-mode EC for H.264/AVC is proposed to reduce the bandwidth of loading SW. Two different strategies aimed at intra-mode and inter-mode reconstructed MB are used to achieve up-to 9.2 times of CR under the lossless-compression constrain. By resource sharing, the corresponding hardware is designed and integrated into H.264/AVC codec with almost no area overhead. The simulation result shows that our EC can reduce 66.2% and 75.3% system bandwidth in average when QPs are 15 and 30, respectively.

Table 4. The overall simulation results of EC (Bamboo:720x480 Wendy:720x576 Crew,Bigships:1280x720).

	CR_{intra}	CR_{inter}	CR_{hybrid}
Bamboo QP=15	5.848	2.433	**4.082**
Bamboo QP=30	25.0	2.933	**4.651**
Wendy QP=15	4.00	1.618	**1.916**
Wendy QP=30	17.86	2.304	**2.786**
Crew QP=15	8.065	3.571	**6.024**
Crew QP=30	55.56	6.173	**9.259**
Bigships QP=15	4.902	2.347	**2.387**
Bigships QP=30	23.81	3.279	**3.279**

7. REFERENCES

[1] Joint Video Team, *Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification*, ITU-T Rec. H.264 and ISO/IEC 14496-10 AVC, May 2003.

[2] *Joint Video Team Reference Software JM8.5*, http://bs.hhi.de/ suehring/tml/download/, 2004.

[3] Peter H. Frencken P.H.N. de with and M. van dar Schaar, "An MPEG decoder with embedded compression for memory reduction," *IEEE Transactions on Consumer Electronics*, vol. 44, no. 3, pp. 545–555, 1998.

[4] M. van der Schaar and P.H.N. de With, "Near-lossless complexity-scalable embedded compression algorithm for cost reduction in DTV receivers," *IEEE Transactions on Consumer Electronics*, vol. 46, no. 4, pp. 923–933, 2000.

[5] Bourge Arnaud and Jung Joel, "Low-power H.264 video decoder with graceful degradation," in *Proc. of Visual Communications and Image Processing*, 2004, vol. 5308, pp. 372–383.

[6] R. v.d. Vleutenl R. Manniesing, R. Kleihorstl and E. Hendrilks, "Implementation of lossless coding for embedded compression," in *Proc. of IEEE ProRISC*, 1998, pp. 385–389.

[7] Andreas E. Savakis, "Evaluation of algorithms for lossless compression of continuous-tone images," *Journal of Electronic Imaging*, vol. 11, pp. 75–86, 2002.

[8] M. Y. Hsu, H. C. Chang, and L. G. Chen, "Scalable module-based architecture for MPEG-4 BMA motion estimation," in *Proc. of ISCAS*, 2001, pp. 245–248.

[9] J. C. Tuan, T. S. Chang, and C. W. Jen, "On the data reuse and memory bandwidth analysis for full-search block-matching VLSI architecture," *IEEE Transactions on CSVT*, vol. 12, pp. 61–72, Jan. 2002.

[10] T. Wiegand, X. Zhang, and B. Girod, "Long-term memory motion-compensated prediction," *IEEE Transactions on CSVT*, vol. 9, pp. 70–84, Feb. 1999.