A 1062Mpixels/s 8192x4320p High Efficiency Video Coding (H.265) Encoder Chip

Sung-Fang Tsai, Chung-Te Li, Hsuan-Hung Chen, Pei-Kuei Tsung, Kuan-Yu Chen, Liang-Gee Chen
Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan

Abstract

A single-chip HEVC (H.265) 8192x4320p encoder is implemented on a 25mm^2 die with 28nm process. It dissipates 708mW at 312MHz for 8192x4320p encoding. Frame-level pipelining reduces 8.90 GB/s external memory bandwidth and improve CABAC rate by 50%. A 7.14MB three-level memory hierarchy is designed to support internal 43.38 GB/s bandwidth and 13-port accesses, and reduces external reference frame bandwidth down to 2.97 GB/s. High complexity mode decision is supported with CFBAC rate estimator.

Keywords: HEVC, UHDTV, video encoder, high complexity mode decision, memory hierarchy, pipeline

Introduction

High Efficiency Video Coding (H.265) is next generation video coding standard, which offers more than 50% bitrate saving over H.264/AVC and enables efficient compression for 4320p Ultra HDTV (Super Hi-Vision) [1]. This paper presents a single-chip HEVC encoder that supports real-time 8192x4320p 30fps coding. The design addresses 3 key challenges: 1) Frame-level data dependencies results in huge external bandwidth; 2) High port number and high bandwidth is required for reference frame access; 3) High complexity mode decision is required for complex prediction modes in HEVC.

To overcome these challenges, several techniques are proposed: 1) Frame-level pipeline reduces 8.90 GB/s external memory bandwidth and improve CABAC rate by 50%. 2) 3-level memory hierarchy provides 13 port concurrent access and internal 43.38 GB/s bandwidth, and reduce external reference frame access to 2.97 GB/s. 3) High complexity mode decision hardware is proposed with low cost context-fixed CFBAC rate estimator.

Fig. 1 shows the block diagram of proposed encoder. A three-stage frame-level scheme pipeline and a 6-stage LCU-level pipeline scheme in the first frame-level stage are adopted. Transform (TF) will refine 6 modes for high complexity mode decision. SAO and ALF are included in this design.

Frame-level Encoder Pipeline Architecture

Fig. 2 shows the proposed frame-level pipeline architecture. The frame-level dependency problem in HEVC [1] is shown in Fig. 2(a). Conventional macro-block (MB) pipeline scheme does not perform well due to several reasons: 1) The new loop filter Sample Adaptive Offset (SAO) and Adaptive Loop Filter (ALF) is frame-level processing; 2) CABAC requires to encode the parameters used in SAO and ALF; 3) For high CABAC throughput, frame must be separated into slices. There is frame-level latency between prediction engine and CABAC engine.

To solve this problem, we use a three-stage frame-level pipelining as Fig. 2(b). We divide the SAO and ALF process

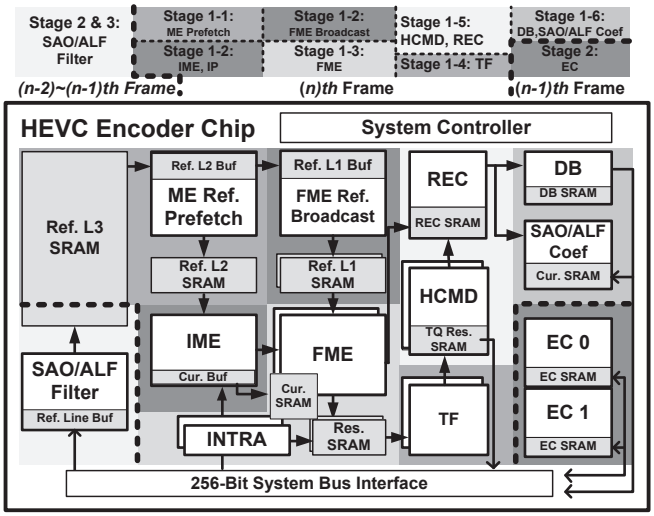

Fig. 1. Block diagram of the proposed HEVC encoder.

into 3 frame-level stages. The first stage is coefficient accumulation, and the others are filtering stage. CABAC is performed at the second stage. In stage 1, the dependency between SAO and ALF coefficient accumulation is removed with 0.005dB quality loss. There are no frame-level dependency anymore and can be performed in LCU pipeline after deblocking (DB). The reconstructed frame is written out after deblocking. In stage 2 and 3, SAO and ALF online filtering is performed once the reference frames are requested by the prediction engine. With online filtering, the additional R/W bandwidth for SAO and ALF filtering is saved. By using this pipelining architecture, 8.90 GB/s bandwidth is saved. No additional reconstructed frame bandwidth is required for SAO and ALF. In addition, the SAO and ALF parameters are available for CABAC encoder in stage 2. To support high bit rate applications, high throughput is achieved by doubling CABAC in stage 2.

Three-Level Memory Hierarchy

Fig. 3 shows the three-level memory hierarchy used for reference frame access and corresponding schedule. Design challenges for reference memory are listed below: 1) For throughput requirement, integer and fractional motion estimation (ME) needs to use up to 13 port for access reference frame. The internal bandwidth for reference frame reaches up to 43.38 GB/s; 2) Conventional scheme as level-C search window memory requires as high as 14.83 GB/s. Besides, it can only support 2 ports by using dual port memory. We propose a 7.14MB three-level memory hierarchy for solving these issues. To reduce the external memory access, a L3 reference SRAM is used for buffering. By using the L3 reference SRAM, the total bandwidth is reduced to 2.97 GB/s.

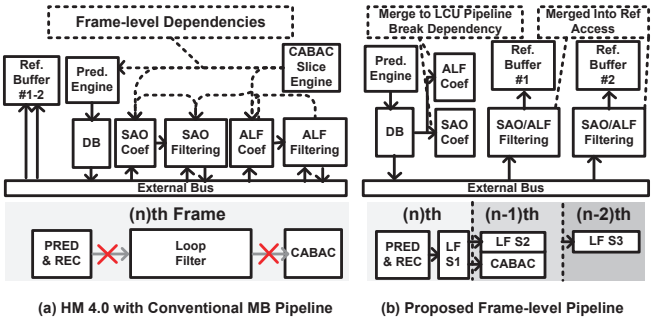

Fig. 2. Frame-level encoder pipeline.

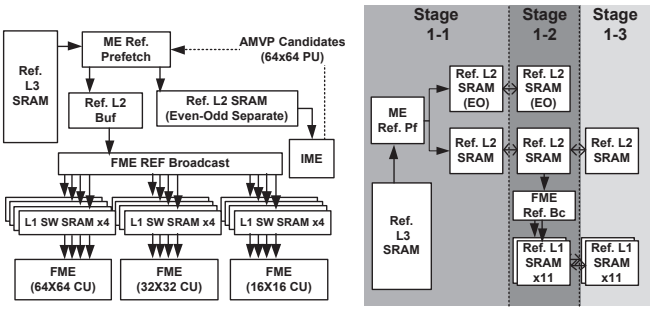

Fig. 3. Three-level memory hierarchy and fetching schedule

To support the 1-port subsampled integer ME and 12-port fractional ME, we use many small L1 and L2 SRAM. Search range is pre-fetched according to the AMVP center. At the next stage, search range is broadcasted to 11 L1 SRAM. Single-port ping-pong style SRAMs are used to forward data to the next stage for the best efficiency.

High Complexity Mode Decision via CFBAC Bit Estimator

Fig. 4 shows the proposed high complexity mode decision (HCMD) scheme. HEVC uses many complex prediction modes that require high complexity rate-distortion optimization for mode decision. For each mode in HCMD, two CABAC modules are required. There are total 17 intra modes and 13 inter modes that need to go through HCMD in HM 4.0. As a result, the overhead is very high. For this issue, two techniques are proposed: 1) Fast decision for the same CU depth is applied. 6 modes are required for HCMD with 0.2dB quality loss. 2) For bit estimation, a CU-based Context-Fixed Binary Arithmetic Counter (CFBAC) is proposed. With fixed context scheme, context memory may be shared among the bit counters. In addition, contexts do not require updating during bits estimation. Bit counts are derived by total sum of ones and zeros and table lookup. With this scheme, 88% gates are saved with only 0.03dB quality loss. This method also outperforms table-lookup only method since table-lookup method still requires context update.

Implementation Results

This chip is implemented on a 25mm^2 die by using TSMC 28nm 1P10M process. The power consumptions are 708mW at 312MHz for 8192x4320p encoding. Fig. 5 shows the chip micrograph and summarizes the chip features. Fig. 6 shows comparisons with previous encoders [2] [3]. Our proposed encoder

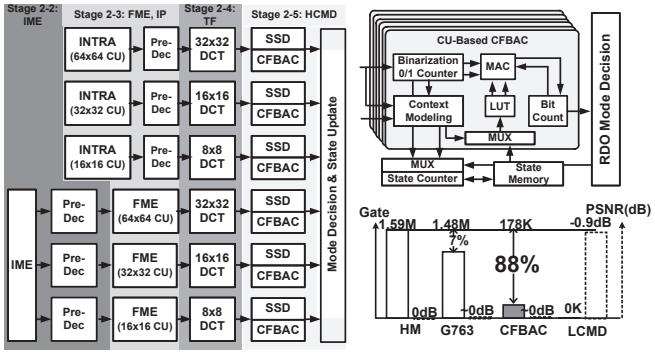

Fig. 4. High complexity mode decision architecture.

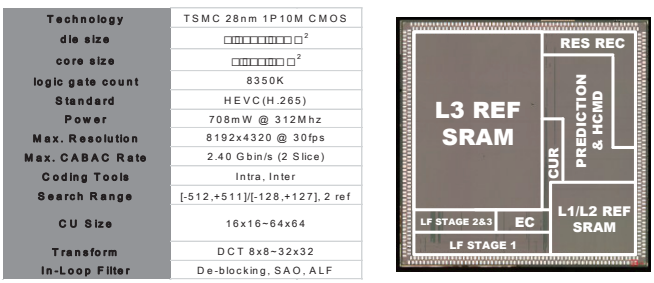

Fig. 5. Chip specification and micrograph.

	ISSCC'09 [2]	VLSIC'12 [3]	This Work
Standard	H.264 High	H.264	HEVC (H.265)
Coding Tools	Intra & Inter	Intra	Intra & Inter
Resolution	4096x2160@24fps	7680x4320@60fps	8192x4320@30fps
Quality (vs JM)	-0.03 dB to -0.08dB	< -3dB	+2.5dB
Max. Throughput	212Mpixels/s	1991Mpixels/s	1062Mpixels/s
Max. CABAC Rate	1.09Gbins/s	1.41Gbin/s	2.40Gbin/s
Technology	TSMC 90nm	e-Shuttle 65nm	TSMC 28nm
Core Size	11.46mm2	2.07mm2	25mm2
Gate Count	1732K	678.8K	8350K
Power	522mW@280Mhz	139.9mW@280Mhz	708mW@312Mhz

Fig. 6. Comparison of coding quality and performance

can have 2.5dB PSNR gain over previous encoders. Supported CABAC rate may up to 2.40Gbin/s by two CABAC slice. By using proposed frame-level pipeline and three-level memory hierarchy, the total bandwidth is reduced below 7GB/s during inter frame encoding at 8192x4320p.

Acknowledgements

The authors thank TSMC University Joint Development Project for technology and process support.

References

[1] Sullivan, G. J.,Ohm, J.-R., Han, W.-J., Wiegand, T., Wiegand, T., "Overview of the High Efficiency Video Coding (HEVC) Standard," *Transaction on Circuits and Systems for Video Technology*, vol.22, no.12, pp.1649-1668, Dec. 2012.

[2] L.-F. Ding, et al., "A 212Mpixels/s 4096x2160p multiview video encoder chip for 3D/quad HDTV applications," *ISSCC Dig. Tech. Paper*, pp. 154-155, Feb., 2009.

[3] Dajiang Zhou, et al., "A 4320p 60fps H.264/AVC intra-frame encoder chip with 1.41Gbins/s CABAC," *VLSI Circuits Symposium*, pp.154-155, Jun. 2012.