

# An H.264/AVC Scalable Extension and High Profile HDTV 1080p Encoder Chip

Yi-Hau Chen<sup>1</sup>, Tzu-Der Chuang<sup>1</sup>, Yu-Jen Chen<sup>1</sup>, Chung-Te Li<sup>1</sup>, Chia-Jung Hsu<sup>2</sup>,  
Shao-Yi Chien<sup>1</sup>, and Liang-Gee Chen<sup>1</sup>

<sup>1</sup>Graduate Institute of Electronics Engineering and Department of Electrical Engineering,  
National Taiwan University, Taipei, Taiwan, <sup>2</sup>UMC, Hsinchu, Taiwan

## Abstract

The first single-chip H.264/AVC HDTV 1080p encoder for scalable extension (SVC) with high profile is implemented on a 16.76mm<sup>2</sup> die with 90nm process. It dissipates 349/439mW at 120/166MHz for high profile and SVC encoding. The proposed frame-parallel architecture halves external memory bandwidth and operating frequency. Moreover, the prediction architecture with inter-layer prediction tools are applied to further save 70% external memory bandwidth and 50% internal memory access.

## Introduction

With the prevalence of streaming multimedia application, the newest H.264/AVC scalable extension, SVC, is set up to provide temporal, spatial and quality scalabilities in single bit-stream [1]. Besides, the H.264/AVC high profile saves 20% to 30% bit rates compared to baseline profile. However, compared to baseline profile [2, 3], high profile requires two times of computation and memory bandwidth, and SVC encoder needs four times of them. In this work, we implement the first HDTV 1080p H.264 encoder for SVC with high profile. The encoding tools such as hierarchical B-frame (HB), inter-layer prediction, context-based adaptive binary arithmetic coding (CABAC), and fine-grain-scalability (FGS) are supported in this chip as defined in reference software JSVM 7.0.

To integrate high profile tools and scalabilities into an HDTV encoder, the several times of computation and memory bandwidth are the main obstacles for system integration. Besides, the temporal and spatial scalability tools lead to different temporal prediction distances and new prediction modes. New data reuse schemes should be developed for various specifications. Finally, to support high bit rate applications (like 10~20Mbps) and FGS, high throughput CABAC and new design methodology of FGS should be explored.

## Frame-Parallel Encoder Architecture

Fig. 1 shows the proposed hardware architecture. The four-stage macro-block (MB) pipeline scheme is adopted with balanced schedule between Integer-pixel Motion Estimation (IME), Fraction-pixel ME (FME), Intra and Reconstruction (REC), and Entropy Coding (EC) and FGS scan engine. The data independency between B-frames is exploited to encode two co-located MBs in two different frames simultaneously, which is called as frame-parallel encoding scheme. For example, two B-frames in IBBP scheme and the two separate B-frames in IBPBP scheme of HB can be concurrently encoded and share the overlapped search window data as shown at the bottom of Fig. 1. With this frame-parallel scheme, the external memory bandwidth of loading search window data

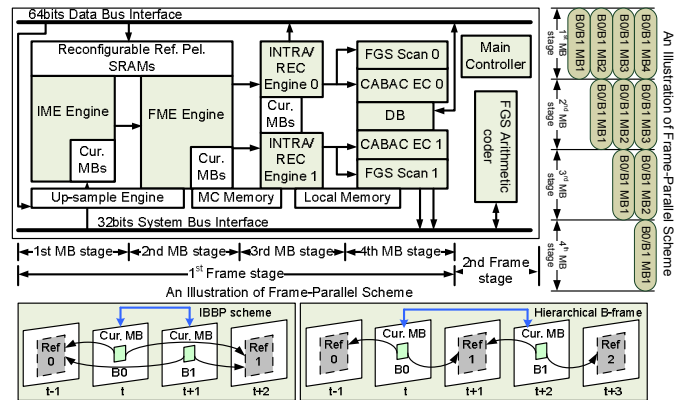


Fig. 1. Block diagram of the proposed H.264 frame-parallel encoder.

can be reduced 50% and 25%, respectively. Moreover, computing cycle for each pipeline stage engine is doubled without increasing operating frequency. It is beneficial for engines of sequential property like EC. In the last two stages, Intra, REC, FGS scan and EC engines are duplicated to process two MBs in the same pipeline stage while de-blocking (DB) engine is shared. The additional frame stage is designed for FGS coding.

## Temporal Prediction with Inter-Layer Prediction

The proposed H.264/AVC scalable extension encoder encodes all supported spatial resolutions, i.e., from QCIF, CIF, 4CIF to 16CIF or HDTV 1080p, and uses inter-layer prediction to reduce the redundancy between spatial layers. Meanwhile, the temporal scalability tool, HB scheme, has several temporal levels, which maps to different temporal distances between current frame and reference frames. Fig. 2 shows the proposed IME architecture. Two main techniques are developed. First, the hierarchical ME algorithm is adopted. It utilizes inter-layer motion prediction to find search predictors. For larger resolution, Level C data reuse is cooperated with Centric Moving Row Buffer (CMRB) derived by averaging up-sampled base-layer motion vectors in the same row and refinement is applied while search predictor is outside CMRB. It can save up to 70% external memory bandwidth with 0.1dB quality loss compared to full search ME algorithm. Second, the reconfigurable search range SRAM has three main buffer arrays and can be configured to support different frame type, temporal distance and frame parallelism as shown at the top of Fig. 2.

Fig. 3 shows proposed FME engine. In addition to our previous proposed centric-quarter scheme [4] which can calculate 25 half/quarter candidates in parallel, two techniques are developed to achieve high throughput and data reuse for inter-layer residual prediction and B-frame refinement. First, the transformed coefficients from normal prediction are reused by residual prediction in transform domain based on Hadamard

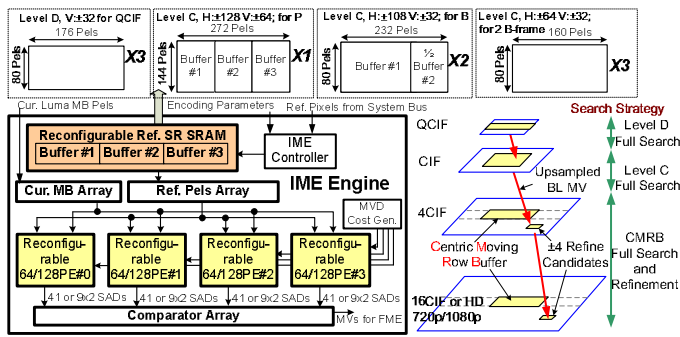


Fig. 2. Algorithm and architecture of IME module.

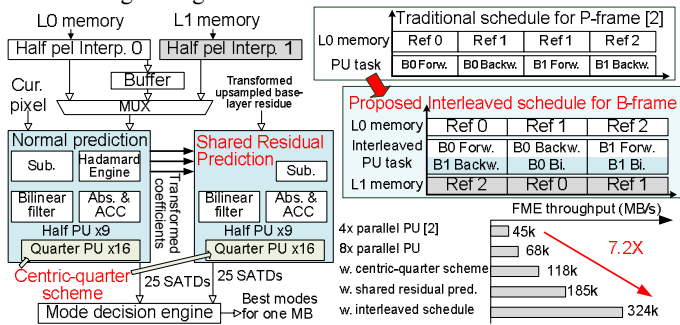


Fig. 3. FME module architecture for B-frame and residual prediction.

transform's linearity. Thus, 50% internal memory bandwidth, pixel interpolation and Hadamard transform can be saved in average. Second, to work with frame-parallel architecture, two interpolation engines are scheduled in an interleaved manner for three reference frames. Two prediction tasks can be processed in the same task stage by buffering interpolated data from L0 memory. Although B-frame and residual prediction will make computation of one MB be six times compared to baseline P-frame, the proposed FME can achieve HDTV 1080p specification with above techniques.

### Multi-symbol CABAC and Low-Bandwidth FGS

Fig. 4 shows the proposed architecture of EC and FGS engine. The arithmetic coder of CABAC is optimized by four-stage pipeline to shorten critical path of each symbol, and an update technique is developed to solve context dependency so four symbols can be processed in one cycle. Thus, 660 Msymbols/sec throughput can be achieved to support HD1080p encoding with 45dB video quality in average. The FGS engine adopts scan bucket method to remove redundant zero coefficients. The early context modeling changes the coefficients into binary contexts. By experiments, 88% external memory bandwidth of FGS can be saved and the generated bitstream can be decoded at any bit rate point for quality scalability.

### Implementation Results

This chip is implemented on a 16.76mm<sup>2</sup> die by using UMC 90nm 1P9M process. The measured power consumptions are 306 and 41mW at 120 and 166MHz for high profile and SVC, respectively. Fig. 5 shows the chip micrograph and Table 1 summarizes the chip features and comparisons with a previous baseline encoder [3]. Fig. 6 shows the coding performance. Our high profile encoder can save 20% to 30% bit rate compared to baseline encoder for "Toy". The scalabilities are also demonstrated by "Rush". Compared to JSVM7.0, the

also demonstrated by "Rush". Compared to JSVM7.0, the quality loss of our chip is less than 0.1dB. The spatial scalability can save 10% to 20% bit rate compared to simulcast.

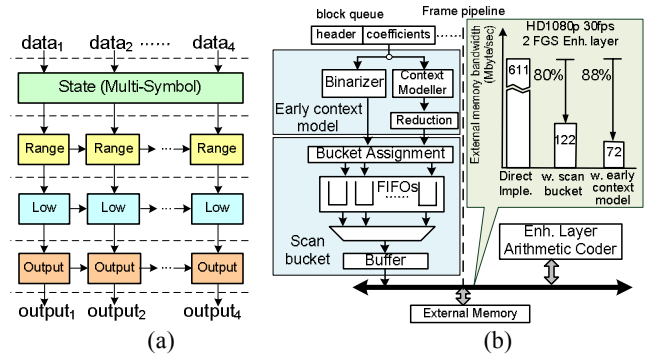


Fig. 4. Block diagram of (a) multi-symbol CABAC and (b) FGS.

	Liu[3]	Proposed
Technology	TSMC0.18um CMOS 1P6M	UMC90nm CMOS 1P9M
Core size	27.1mm <sup>2</sup>	16.76mm <sup>2</sup>
Logic Gates	1140k	2079k
SRAMs	108.3KB	81.7KB
H.264 Profile	Baseline 1080p	Scalable/High*1080p
Ref. Num	1	2 (B-frame)
Slice Type	I, P	I, P, B, EI, EP, EB
Search Range	196x128	512x256/128x64*
Entropy Coder	CAVLC	CABAC+FGS
Frequency	200MHz	166/120*MHz
Power	1409mW	411/306*mW

\*: the spec. for high profile, and the other is for SVC encoding.

Table 1: Performance Comparison

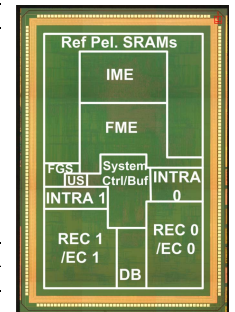


Fig. 5. Chip micrograph

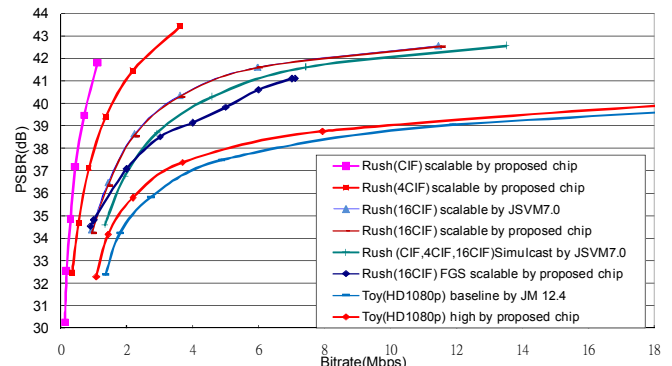


Fig. 6. Comparison of coding quality and scalability

### Acknowledgements

Authors thank UMC University Program team and process support.

### References

- [1] Heiko Schwarz, et al., "Overview of the scalable video coding extension of the H.264/AVC standard," *Transaction on circuits and systems for video technology*, pp.1103-1120, Sep. 2007.
- [2] Y.-W. Huang, et al., "A 1.3TOPS H.264/AVC Single-Chip Encoder for HDTV Applications," *ISSCC Dig. Tech. Paper*, pp. 128-129, Feb., 2005.
- [3] Z. Liu, et al., "A 1.41W H.264/AVC real-time encoder SOC for HDTV1080P," *VLSI Circuits Symposium*, pp.12-13, Jun. 2007.
- [4] T.-C. Chen, et al., "2.8 to 67.2mW low-power and power-aware H.264 encoder for mobile applications," *VLSI Circuits Symposium*, pp.22 2-223, Jun. 2007.