

## HARDWARE ARCHITECTURE DESIGN FOR VISUAL PROCESSING: PRESENT AND FUTURE

*Po-Chih Tseng and Liang-Gee Chen*

DSP/IC Design Lab, Graduate Institute of Electronics Engineering,  
Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan  
E-Mail: {pctsensg, lchen}@video.ee.ntu.edu.tw

### ABSTRACT

This paper presents the present and future trends of hardware architecture design for image and video coding. Fundamental design issues are discussed with particular emphasis on efficient dedicated implementation. Hardware architectures for MPEG-4 video coding and JPEG 2000 still image coding are reviewed as examples, and special approaches exploited to improve efficiency are identified. Further perspectives are also presented to address the challenges of hardware architecture design for future image and video coding.

### 1. INTRODUCTION

Due to advances in image and video coding algorithms as well as in very large scale integration (VLSI) technology, diverse and interesting visual experiences have been brought to human daily life. A number of international standards have contributed to the great success of image and video coding applications. Still image compression applications, such as digital still cameras, are covered by the ISO/IEC JPEG standards, both the current JPEG and the emerging JPEG 2000. The present MPEG-1 and MPEG-2 of ISO/IEC MPEG standards are used for video storage and playback, digital TV broadcast, and video-on-demand applications, while the emerging MPEG-4 intends to cover wide-ranging multimedia communications applications. Video communications applications, such as video telephony and video conference, are regulated by the ITU-T H.26X standards, including early-stage H.261, present H.263, and the new generation H.264/AVC.

The availability of low-cost and low-power hardware with sufficiently high performance is essential for the popularization of image and video coding applications. Thus, efficient hardware implementations in VLSI are of vital importance. However, image and video coding algorithms are characterized by very high computational complexity. Real-time processing of multi-dimensional image and video signal involves operating continuous data streams of huge volumes. Such critical demands cannot be fulfilled by conventional hardware architectures without specific adaptation. Therefore, special architectural approaches are indispensable for efficient hardware solutions to meet real-time constraints with desired low cost and low power.

Since emerging MPEG-4 and JPEG 2000 are capable of offering both improved coding efficiency and additional functionalities beyond present standards, these advanced features further increase the computational complexity and consequently pose challenges for hardware architecture design. This paper presents the present and future trends of hardware architecture design for image

and video coding, primarily focusing on emerging MPEG-4 and JPEG 2000. An overview of fundamental design issues is given in Section II. Hardware architectures for MPEG-4 video coding and JPEG 2000 still image coding are reviewed as examples in Sections III and IV, respectively. Finally, the further perspectives in Section V are presented to address the challenges of hardware architecture design for future image and video coding.

### 2. FUNDAMENTALS OF HARDWARE ARCHITECTURES

Efficient dedicated hardware implementations for image and video coding rely on the thorough analysis of target algorithms and the exploitation of special computational characteristics inherent in the algorithms. The design goal is to achieve a dedicated system architecture with the highest degree of adaptation. The design approach is to perform the mapping of individual tasks onto different module architectures, and then to optimize each module architecture in terms of performance, area, and power constraints.

In addition to module architectures, a complete system architecture also includes memory architecture and interconnect architecture. As image and video coding algorithms involve a large amount of data computation, during processing continuous data streams in huge volumes, a large amount of data communication is also incurred in two aspects. The first is the data access through the frame buffer, which is mainly dominated by the memory architecture; and the other is the data access between different module architectures and the memory architecture, which is mainly dominated by the interconnect architecture. Data communication has already become a bottleneck for complete system architectures and has a significant impact on the overall system performance and power consumption. Below, special architectural approaches with regard to memory architecture and interconnect architecture are discussed.

Data access through the frame buffer is a slow and power-consuming process, and two architectural approaches can be applied to relieve this problem. The first approach is to adopt special local memory buffers. Because data access patterns of most tasks of image and video coding are predictable, and there are also

major concern for integration. Due to the advances in VLSI technology, various designs have already integrated embedded DRAM or SRAM as on-chip frame buffer to solve the data access problem.

In addition to memory architecture, the interconnect architecture that communicates data between different module architectures and the memory architecture is another key factor for data communication. In-depth understanding of the inter-module communication patterns is essential to design an efficient interconnect architecture, which can provide higher bandwidth and consume lower power. There is a trade-off between flexibility and efficiency of interconnect architecture. For example, the global bus provides higher flexibility but lower efficiency, whereas the dedicated data link with full adaptation to specific algorithm provides the highest efficiency but without flexibility.

### 3. HARDWARE ARCHITECTURES FOR MPEG-4

In this section, several of the most representative designs for MPEG-4 are selected as examples [1, 2, 3, 4, 5, 6, 7]. The detailed system architectures, including module architectures, memory architecture, and interconnect architecture, are discussed below, and comparisons of these designs are also made.

Takahashi et al. [1] present the first MPEG-4 video codec design, as shown in Fig. 1. Several dedicated module architectures are adopted for computation-intensive ME/MC, DCT/IDCT, and VLC/VLD, while an embedded RISC processor is included to provide flexibility for other tasks. Special local memory (LM) buffers are used to reduce the data access through off-chip frame buffer. A DMA controller that connects any two functional blocks acts as the interconnect architecture to provide higher power efficiency than a global bus architecture. The ME adopts a fast search algorithm with search range of  $-15.5/+31.5$ . This chip consumes 60mW at 30MHz for simple profile QCIF 10 frames/sec encoding and decoding.

Based on [1], Nishikawa et al. [2] present a single-chip MPEG-4 video phone design with embedded DRAM. In addition to video codec, a speech codec, a multiplexer, and several I/O units are also included. This chip consumes 240mW at 60MHz for full system functionality, including simple profile QCIF 15 frames/sec encoding and decoding of the video codec part. The integration of embedded DRAM significantly reduces a large amount of I/O power consumption caused by the data access through frame buffer.

Another single-chip MPEG-4 audiovisual design based on previous two designs [1] and [2] is presented by Arakida et al. [6]. In addition to original modules of previous designs, a 5GOPS adaptive filter engine is also included for post-processing. With more advanced VLSI technology as well as several low-power design techniques, this chip consumes 160mW at 125MHz for full system functionality, including simple profile CIF 15 frames/sec encoding of the video codec part.

Hashimoto et al. [3] present the first MPEG-4 video codec design with the support of core profile. Similar to [1], several dedicated module architectures are adopted for computation-intensive ME/MC, DCT/IDCT, VLC/VLD, padding, and shape decoding, while a programmable DSP is used for other tasks. The embedded DRAM is integrated to reduce the I/O power consumption, and special local memory buffers are adopted to reduce the data access through frame buffer. A global bus is used as the interconnect architecture for higher flexibility. This chip consumes 90mW at 54MHz for simple profile QCIF 15 frames/sec encoding and decoding or core profile CIF 15 frames/sec decoding.

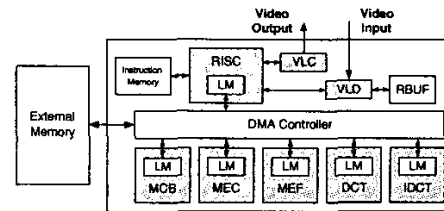


Fig. 1. Toshiba MPEG-4 video codec architecture.

Based on [3], Ohashi et al. [4] present a low-power MPEG-4 video decoder design. Rather than embedded DRAM, the embedded SRAM is integrated as the frame buffer for ease of future system-on-a-chip integration. Due to the elimination of unnecessary modules from previous design [3], this chip consumes only 11.1mW at 27MHz/54MHz for simple profile QCIF 15 frames/sec decoding.

A highly efficient MPEG-4 video codec design with full adaptation approach is proposed by Nakayama et al. [5], as shown in Fig. 2. Dedicated module architectures are adopted for all coding tasks including codec control. In addition, the dedicated data link is used as the interconnect architecture for maximum data communication efficiency. The data access and the size of local memory buffers are reduced due to the adoption of dedicated data link. The ME adopts a fast scene-adaptive search algorithm with search range of  $-15.5/+15.5$ . As a consequence of full adaptation, this chip consumes only 29mW at 13.5MHz for simple profile CIF 15 frames/sec encoding and decoding.

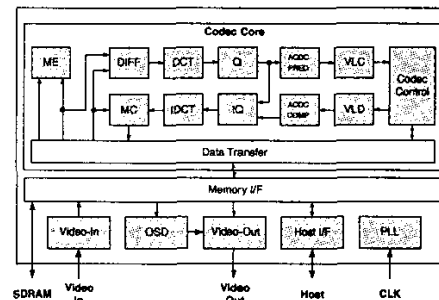


Fig. 2. Fujitsu MPEG-4 video codec architecture.

Completely different from previous designs, Stolberg et al. [7] present a fully programmable multicore system-on-a-chip, comprised of a 16-way SIMD DSP core with a 2-D matrix memory, a 64-bit VLIW DSP core with subword parallelism, and a 32-bit RISC core. These programmable cores are connected with a global bus, and special local memory buffers are adopted to communicate data between each other. This chip consumes 3.5W at 145MHz for advanced simple profile D1 25 frames/sec decoding or simple profile D1 25 frames/sec encoding.

Table 1 shows the comparison of several architecture parameters of the aforementioned system architectures, and Table 2 shows the comparison of several chip parameters of these two architectures. From these tables, three trends can be summarized as follows. First, except for the fully programmable solution [7], the

specification of MPEG-4 video codec designs has increased year by year with reasonable power consumption. Second, the integration of embedded DRAM or SRAM as the frame buffer has become the mainstream. Finally, fast search algorithms for ME are adopted by most designs due to stringent power constraints.

**Table 1.** Comparison of MPEG-4 architecture parameters.

Architecture	Category	Function	Profile Level	ME Search Algorithm	Frame Buffer
Takahashi [1]	Dedicated	V Codec	< SP@L1	Fast Search	External
Nishikawa [2]	Dedicated	AVS Codec	SP@L1	Fast Search	eDRAM
Hashimoto [3]	Dedicated	V Codec	CP@L1	N/A	eDRAM
Ohashi [4]	Dedicated	V Decoder	SP@L1	No	eSRAM
Nakayama [5]	Dedicated	V Codec	SP@L3	Fast Search	External
Arakida [6]	Dedicated	AVS Codec	SP@L2	Fast Search	eDRAM
Stolberg [7]	Programmable	V Codec	ASP@L5	Fast Search	External

Note: In Function part, V means Video and AVS means Audio-Video-System

**Table 2.** Comparison of MPEG-4 chip parameters.

Architecture	Technology ( $\mu\text{m}$ )	Frequency (MHz)	Power (mW)	Area ( $\text{mm}^2$ )	Video Encoding Capability
Takahashi [1]	0.3	30	60	81*	QCIF@10fps
Nishikawa [2]	0.25	60	240	117	QCIF@15fps
Hashimoto [3]	0.18	54	90	75	QCIF@15fps
Ohashi [4]	0.18	27/54	11.1	37	No
Nakayama [5]	0.18	13.5	29	28*	CIF@15fps
Arakida [6]	0.13	125	160	43	CIF@15fps
Stolberg [7]	0.18	145	3500	81*	D1@25fps

Note for \*: The area does not include frame buffer

#### 4. HARDWARE ARCHITECTURES FOR JPEG 2000

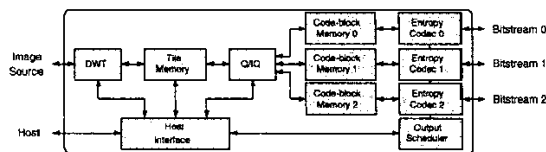
In this section, several system architectures for JPEG 2000 are chosen as examples. Some of them are complete system architectures, including DSPworx [8], Yamauchi [9], Andra [10], ADI [11], and Fang [12], while some of them are hardware accelerators only and need to cooperate with a host processor, such as ALMA [13] and AMPHION [14].

In order to obtain highly efficient hardware solutions, all of the system architectures adopt several dedicated module architectures for computation-intensive tasks such as DWT, EBCOT Tier-1, and quantization. The complete system architectures [9, 8, 10, 11, 12] also include dedicated module architectures for EBCOT Tier-2 with rate control. As for memory architecture, all system architectures heavily rely on various special local memory buffers in order to effectively reduce the large amount of data access through off-chip frame buffer. In addition, the dedicated data link is adopted as the interconnect architecture for all system architectures to achieve highly efficient data communication. Below, three selected designs are discussed in detail, and the comparisons of architecture parameters and chip parameters are also made.

Amphion [14] reports a JPEG 2000 codec hardware accelerator design, as shown in Fig. 3. In order to increase the throughput, three entropy codecs are adopted to encode or decode three code blocks in parallel. This design can achieve 60/20 Msamples/sec encoding/decoding rate at 180MHz operating frequency.

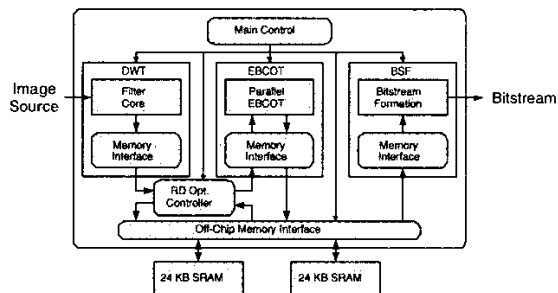
A complete JPEG 2000 encoder system-on-a-chip is proposed by Yamauchi et al. [9]. It uses block-based 2-D DWT

to minimize the size of local memory buffers for DWT, but comes with increased data access. For EBCOT Tier-1, it encodes two bit-planes, three coding passes, and four symbols in parallel to increase the throughput by a factor of 24 compared with the conventional approach. As a result of these techniques, this chip can encode 20.7 Msamples/sec at 27MHz. In addition, a one-pass code size controlling method is used to predict the code size for rate control.



**Fig. 3.** Amphion JPEG 2000 codec hardware accelerator.

Based on parallel EBCOT Tier-1 architecture, Fang et al. [12] propose a JPEG 2000 encoder architecture, as shown in Fig. 4. Unlike conventional architectures, this architecture does not require any code block memory due to the use of parallel EBCOT Tier-1 architecture. This design is fully pipelined, and the throughput rate is equal to the operating frequency. Another important feature of this design is the Rate-Distortion Optimized (RD Opt.) controller for rate control, which realizes the pre-compression rate-distortion optimization algorithm. By using this algorithm, redundant computation and data access, as well as bitstream buffer are eliminated, which leads to low power and small area.



**Fig. 4.** Fang's JPEG 2000 encoder architecture.

Table 3 shows the comparison of several architecture parameters of all system architectures. Among these parameters, the tile size may be the most important, since it affects the architectures for DWT. Line-based 2-D DWT can minimize the off-chip data access, but it requires a large-size on-chip memory, which is proportional to the tile size. On the other hand, the memory size is irrelevant to tile size for block-based 2-D DWT. Thus, block-based 2-D DWT is preferred for designs with large tile size. The DWT filter type and decomposition level as well as EBCOT Tier-1 code block size have some effects on coding efficiency. The 9/7 filter requires more processing elements than the 5/3 filter, while higher decomposition level and larger code block size lead to larger memory size.

Table 4 compares several chip parameters of actually implemented system architectures. It is readily seen that the throughput of Fang's architecture is equal to the operating frequency. More-

over, the area of Fang's architecture is also the smallest since it uses parallel EBCOT Tier-1 architecture. In other architectures, parallel processing of code blocks, which leads to large area, is used to increase the throughput. Another way to increase the throughput is to raise the operating frequency. However, this would result in higher power consumption.

**Table 3.** Comparison of JPEG 2000 architecture parameters.

Architecture	Function	DWT Filter	DWT Level	Tile Size	CB Size
DSPworx [8]	Codec	5/3 + 9/7	5	256x256	N/A
ALMA [13]	Encoder*	5/3 + 9/7	N/A	256x256	64x64
AMPHION [14]	Codec*	5/3 + 9/7	5	128x128	N/A
Yamauchi [9]	Encoder	5/3 + 9/7	2	1024x512	32x32
Andra [10]	Encoder	5/3 + 9/7	5	128x128	32x32
ADI [11]	Codec	5/3 + 9/7	6	4096x2048	N/A
Fang [12]	Encoder	5/3	2	128x128	64x64

Note for \*: The design is a hardware accelerator only but not a complete system architecture

**Table 4.** Comparison of JPEG 2000 chip parameters.

Architecture	Technology ( $\mu\text{m}$ )	Area ( $\text{mm}^2$ )	Frequency (MHz)	Throughput (MS/s)
DSPworx [8]	0.18	289	200	50
AMPHION [14]	0.18	5.4	120	60
Yamauchi [9]	0.25	13.7*	27.4	21
ADI [11]	0.18	144	N/A	65
Fang [12]	0.25	5.5*	81	81

Note for \*: The size is the core size but not die size

## 5. FURTHER PERSPECTIVES

Although the emerging MPEG-4 and JPEG 2000 are capable of providing improved coding efficiency and additional functionalities, the demands on much better coding efficiency and much richer functionalities for image and video coding applications are ever-increasing. Therefore, advanced coding algorithms are continuing to be developed vigorously, and the standardization of new generation coding standards will keep emerging. In order to provide better coding efficiency and richer functionalities, more complicated coding tools must be adopted by advanced coding algorithms, which will further inevitably increase the computational complexity. Even if the advances in VLSI technology continue to provide more processing capability without increased hardware cost and power consumption, there is still a strong need to explore highly efficient hardware architectures for future image and video coding.

For example, the new generation H.264/AVC video coding standard provides a significant improvement in coding efficiency compared with various preceding standards. It can achieve essentially the same reproduction quality as previous standards, while typically requiring 60% or less of the bit-rate. However, the improvement in coding efficiency by H.264/AVC comes, not surprising, with a high degree of computational complexity. An estimation reports that the complexity of H.264/AVC encoder grows an order of magnitude higher than that of the MPEG-4 encoder [15]. Even with highly optimized encoder implementation, the complexity of H.264/AVC is still about 3.4 times more than that of H.263 [16]. Therefore, this advanced video coding standard has a significant impact on computational complexity, and it will consequently present new challenges for hardware architecture design. Special

architectural approaches introduced by this paper are also applicable to future image and video coding. Exploiting these approaches in a more efficient way is a critical factor to cope with the new challenges.

## 6. REFERENCES

- [1] M. Takahashi et al., "A 60 mW MPEG4 video codec using clustered voltage scaling with variable supply-voltage scheme," in *Digest of Technical Papers of IEEE International Solid-State Circuits Conference*, 1998.
- [2] T. Nishikawa et al., "A 60 MHz 240 mW MPEG-4 video-phone LSI with 16 Mb embedded DRAM," in *Digest of Technical Papers of IEEE International Solid-State Circuits Conference*, 2000.
- [3] T. Hashimoto et al., "A 90 mW MPEG4 video codec LSI with the capability for core profile," in *Digest of Technical Papers of IEEE International Solid-State Circuits Conference*, 2001.
- [4] M. Ohashi et al., "A 27 MHz 11.1 mW MPEG-4 video decoder LSI for mobile application," in *Digest of Technical Papers of IEEE International Solid-State Circuits Conference*, 2002.
- [5] H. Nakayama et al., "An MPEG-4 video LSI with an error-resilient codec core based on a fast motion estimation algorithm," in *Digest of Technical Papers of IEEE International Solid-State Circuits Conference*, 2002.
- [6] H. Arakida et al., "A 160mW, 80nA standby, MPEG-4 audiovisual LSI with 16mb embedded DRAM and a 5GOPS adaptive post filter," in *Digest of Technical Papers of IEEE International Solid-State Circuits Conference*, 2003.
- [7] H.-J. Stolberg et al., "An SoC with Two Multimedia DSPs and a RISC Core for Video Compression and Surveillance," in *Digest of Technical Papers of IEEE International Solid-State Circuits Conference*, 2004.
- [8] DSPworx Cheetah, <http://www.dspworx.com/cheetah.htm>.
- [9] H. Yamauchi et al., "Image processor capable of block-noise-free JPEG2000 compression with 30frames/s for digital camera applications," in *Digest of Technical Papers of IEEE International Solid-State Circuits Conference*, 2003.
- [10] K. Andra, C. Chakrabarti, and T. Acharya, "A high-performance JPEG2000 architecture," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 3, pp. 209-218, Mar. 2003.
- [11] Analog Devices Inc. ADV202, <http://www.analog.com/>.
- [12] H. C. Fang et al., "81 MS/s JPEG 2000 single-chip encoder with rate-distortion optimization," in *Digest of Technical Papers of IEEE International Solid-State Circuits Conference*, 2004.
- [13] ALMA Technologies JPEG2KE, <http://www.alma-tech.com/>.
- [14] AMPHION CS6590, <http://www.amphion.com/cs6590.html>.
- [15] K. Denolf, C. Blanch, G. Lafruit, and A. Bormans, "Initial memory complexity analysis of the avc codec," in *Proc. of IEEE Workshop on Signal Processing Systems*, 2002.
- [16] V. Lappalainen, A. Hallapuro, and T. D. Hamalainen, "Performance of H.26L video encoder on general-purpose processor," *Journal of VLSI Signal Processing*, vol. 34, no. 3, pp. 239-249, July 2003.