Determine key profiles for video playback

By Jonah Probell
Technical Marketing Manager
Mobile Multimedia Products
Tensilica Inc.

In the near future, the modern family's living room will be equipped with a DTV with an IPTV and HDTV STB and a high-definition video disk player. Additionally, mobile devices are increasingly adding playback function for entertainment media. Soon all devices will play professionally-produced entertainment content.

Designers of next-generation home and portable devices with video support must understand what is important to consumers as they work on their next SoC design architectures. There are now a number of IP vendors offering video processor cores or hardware blocks that can be designed into SoCs. When evaluating these offerings, it is important to understand how the market is evolving to support different video standards.

Various video standards

Although H.264 has emerged as the preeminent video coding standard for future systems and devices, there are more standards in wider use than ever before. Next-generation video disk and DTV standards also call for Windows Media Video 9/SMPTE VC-1. Some IPTV systems use MPEG-4 and its cousins. MPEG-2 is required for legacy purposes in many systems. H.263 remains widely used in real-time full-duplex video conferencing systems. AVS is the primary coding standard for China's national broadcast purposes.

Different types of devices have different requirements. For example, an STB must decode high-quality video while a video conferencing device must simultaneously decode and encode video in real-time. Battery-powered devices require low-energy consumption, and devices with mobile antennas require particularly strong bit stream error resilience.

The world of video processing is even more complex because many coding standards are divided into profiles that use coding techniques beneficial to different applications and devices. In general, profiles for two-way communication require real-time encoding and low-complexity error resilience while profiles for professionally encoded entertainment video require higher compression quality and low cost of decode. The table shows both the real-time and the entertainment profiles of major compression standards.

It will be crucial for future video devices, even mobile ones, to decode the entertainment video profiles such as H.264 Main profile. For example, a fully capable DVB-H decoder in a handheld device must include support for both H.264 Baseline and Main profile. An end-user product must include the right decoder profiles to be successful in the marketplace.

Buyers of the early RCA Lyra portable media player were frustrated when they discovered that their players could not play any video coded with B frames. While B frames improve compression quality, they are difficult to encode in real-time so they are only used as part of entertainment profiles. This deficiency of the Lyra player left most videos downloaded from the Internet unusable. Such profile deficiencies lead to the quick death of a video player product line in increasingly competitive consumer-product markets. In the case of the Lyra, it left the door open for Archos' stellar success.

Multiformat programmable video processors are required to handle the diversity of coding standards and profiles. The desire by consumers to be entertained as well as to communicate drives a need by chipmakers to use a programmable video processor that can handle entertainment as well as real-time profiles. Conventional wisdom is that hard-wired video blocks are smaller in area, but when different blocks are needed for different standards, the video function can become much larger than when done in a processor. Also, hard-wired video blocks are inflexible to handle changes in these evolving video standards. Therefore, most chip designers today will only consider a programmable video processor.

Efficient handling

However, just because a processor is programmable doesn't mean it can efficiently handle every standard. General-purpose embedded CPUs are not well equipped to handle video streams and would have to run at extremely high megahertz just to decode a low-quality video. Such an approach is not power-efficient for

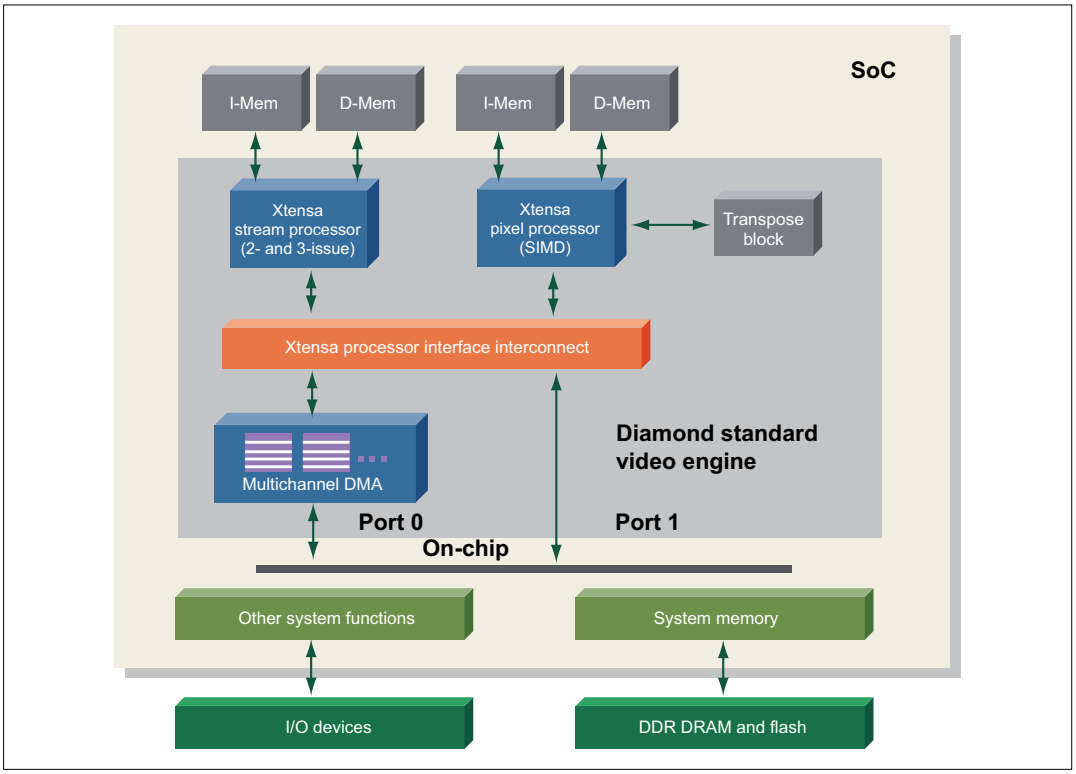

Figure 1: This design includes two Tensilica Xtensa configurable processors and a DMA controller.

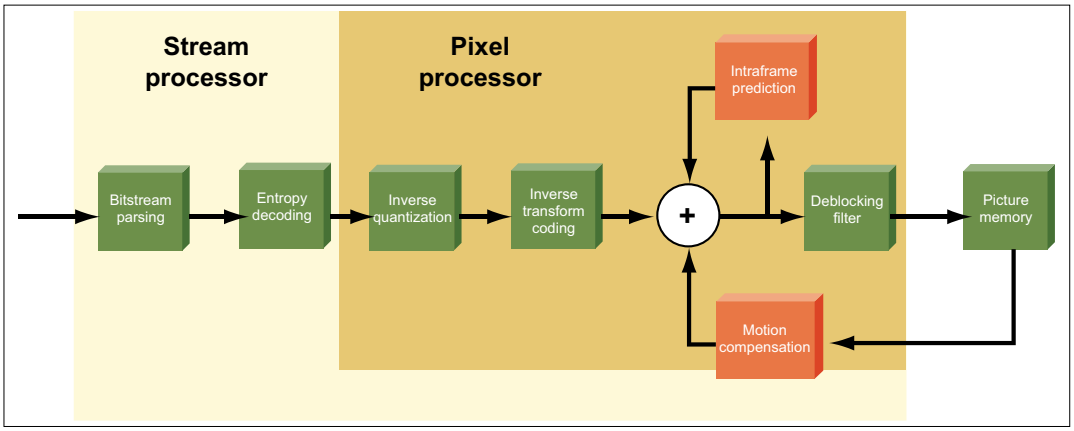

Figure 2: The Stream Processor performs bitstream parsing and entropy decoding.

portable devices. Instead, specialized, video-specific processors are required with special instructions that enable, for example, parallel processing of pixel data using single instruction, multiple data (SIMD) techniques or special instructions for serial processing of video data (entropy decoding, motion-vector prediction etc.).

It is especially difficult to implement processors for H.264 Main profile. H.264 Main profile uses a Context Adaptive Binary Arithmetic Coding (CABAC) method of lossless entropy coding of the bitstream. To decode each binarized element (known as a bin) from the CABAC bitstream depends on the result of the fully decoded previous bin, and each can have a major impact on the status of the decoder. Efficient performance during high bit rate operation is a proven result of such an approach.

By implementing the full entropy decoding process in software using instruction set extensions, Tensilica was able to create a low MHz and power efficient video processor capable of handling difficult bit streams. Tensilica was able to implement full D1 Main profile decoding on a 5Mbps stream at just 162MHz.

Similarly, H.264 Main profile includes support for B frames and interlaced video content, which present difficult challenges for inexperienced developers of video processors and codecs. The simplest solution to difficult video codec problems is to throw more DRAM bandwidth at the problem. That works well in high-end PCs, but is impractical for embedded systems. Because of power consumption and cost constraints, embedded systems cannot tolerate wasted DRAM bandwidth.

Figure 1 shows a block diagram of Tensilica Inc.'s Diamond Standard 388VDO Video Engine. This design includes two Tensilica Xtensa configurable processors and a DMA controller to take advantage of the parallelism inherent in video-compression and video-decompression algorithms. The Stream and Pixel Processors inside the Diamond 388VDO core split the video-compression tasks while the DMA controller moves uncompressed and compressed images into and out of the core and between the two processors. Each processor has its own local instruction and data RAMs.

The Diamond Video Engine's Stream and Pixel Processors are based on Tensilica's configurable Xtensa processor architecture. The Stream Processor has been augmented with additional instructions to perform bitstream parsing and entropy coding. Some of these new instructions are based on Tensilica's flexible-length instruction extensions and employ a VLIW format with two independent operations per instruction. The processor has been augmented with SIMD instructions that perform operations on multiple pixels simultaneously.

The instructions added to both processors allow the engine to encode MPEG-4 Advanced Simple Profile bitstreams and to decode H.264/AVC Main Profile, MPEG-4 Advanced Simple Profile, MPEG-2 Main Profile, and VC-1/WMV 9 Main Profile video bitstreams at standard-definition (or D1) display resolution and 30fps while running at clock rates below 200MHz. Low clock rates generally mean lower power requirements, and the 200MHz clock-rate target was selected so that the Diamond Video Engine could be implemented in a generic, low-cost 130nm IC-fabrication process.

Figure 2 shows the task allocation within the Diamond Video Engine while decoding H.264/AVC video bitstreams. The Stream Processor performs bitstream parsing (separating the Network Abstraction Layer, the Picture Layer, and the Slice Layer) and entropy decoding. The Pixel Processor performs inverse quantization, inverse transform coding, intra-frame prediction, motion compensation, and image deblocking. The Stream Processor assists the Pixel Processor with motion compensation.

Note that it would have been possible to run all of these decoding tasks on one processor, but at a much higher clock rate that would have required a more expensive process technology. Due to the need to minimize power dissipation in portable, battery-powered video products, the Diamond 388VDO Video Engine minimizes power dissipation by keeping clock rates low even when decoding standard-definition video.

Video decode task

When evaluating different video processors, it's important to check to make sure the cores offload the full video decode task—including all bit-stream parsing—from the system host CPU. Conventional video coprocessors only offload the pixel processing functions like motion estimation, and leave a large compute burden on the system controller. This overhead could require that the SoC design use a larger, more power-hungry system controller, an expensive decision in most portable devices.

Studying the leading embedded video devices appearing in homes and hands around the world, it becomes apparent that the most suitable video processors are ones that handle entertainment profiles at high bit rates and do so with low memory bandwidth requirements. With a good understanding of the different video profiles and a knowledge of which profiles need to be implemented in next-generation devices, SoC architects can be well prepared to evaluate the options offered by different IP vendors.

Compression standard	Real-time profile	Entertainment profile
H.264 (AVC)	Baseline	Main
AVS	Part 7 / AVS-M /Jiben profile	Part 2 /Jizhun profile
WMV9 / VC-1	Simple	Advanced
MPEG-4	Simple	Advanced Simple
MPEG-2	N/A	Main
H.263	Profile 0	N/A

Table: Profiles implement different parts of compression standards.