Decode MPEG-2 Video with Virtex FPGAs

Amphion's CS6651 video decoder enables the decompression of video streams in real time.

by Rick Richmond
Amphion
rick@amphion.com

MPEG-2 is the digital video paradigm of today. It is at the heart of the Digital Video Broadcast (DVB) and Advanced Television Systems Committee (ATSC) standard and high-definition digital television systems and DVD-video, which has seen incredible market growth in recent years.

The widespread adoption of these applications and systems, coupled with considerable investment by broadcasters and distributors, indicate that MPEG-2 is going to be around for a good while to come – despite the emergence of new, even better video compression algorithms. Future consumer digital applications, in which audio, video, and data networking technologies converge, are certain to need built-in MPEG-2 video capability.

Yet MPEG-2 video is fairly complex and computationally intensive to decode. The main features of the algorithm are discrete cosine transform (DCT)-based compression and motion estimation techniques. Until now, decoder implementations for digital STBs (Set Top Boxes) and DVD-video players had been the domain of ASIC implementations or software running on very powerful processors.

As the sophistication of products like STBs grows, designs will require ever-increasing flexibility and ever-decreasing development time scales. Now, using the Xilinx Virtex™ series of FPGAs and a new intellectual property (IP) core from Amphion, building MPEG-2 video into your designs is simple. Amphion's CS6651 video decoder, which incorporates an integrated external SDRAM memory controller and display direct memory access (DMA), is a great example of the Platform FPGA capability of the Xilinx Virtex series. This solution allows decoding of MP@ML MPEG-2 video with NTSC or PAL frame rates and resolutions.

The MPEG-2 Video Algorithm

MPEG-2 MP@ML video provides a generic video compression solution for applications such as satellite, terrestrial, and cable

television (in DVB and ATSC formats), as well as optical storage (DVD-video) at NTSC, PAL, and SECAM resolution and frame rates.

MPEG-2 video sequences are composed of three different types of pictures:

- Intra coded pictures (I-pictures), which are compressed using DCT-based techniques

- Predictive coded pictures (P-pictures), which use motion compensation to predict the current picture from a past reference picture

- Bidirectional predictive coded pictures (B-pictures), which use predictions from past and future reference pictures. B-type pictures are not themselves used as reference pictures.

DCT-based Compression

In MPEG-2 MP@ML video, each picture is broken down into 16x16-sized blocks of luminance samples, called macro blocks. These blocks are further divided into 8x8 blocks. Each macro block has six blocks in total: four luminance blocks and two sub sampled chrominance blocks (a 4:2:0 chrominance format). In the encoding process, a two-dimensional, 8-point DCT is applied to each 8x8 block; the resulting coefficients are quantized using a 64-element quantization matrix.

This process reduces amplitude and increases the number of zero-value coefficients. The quantized DCT coefficients are reordered in a zigzag fashion into a one-dimensional stream, effectively grouping together runs of zero-valued coefficients interspersed with non-zeros. This stream of run-level pairs is then encoded using Huffman-style variable length codes (VLCs) based on a statistical model.

Motion Compensation

In predictive and bidirectionally predictive coded pictures, each macro block may have a number of pairs of motion vectors. These specify the horizontal and vertical displace-ment from the current position at which the stored reference picture best resembles the current macro block. The difference (if any) between the motion-compensated prediction and the actual image is coded using the DCT-based techniques described above. The sum of the predicted samples and the prediction error give the final reconstructed macro block.
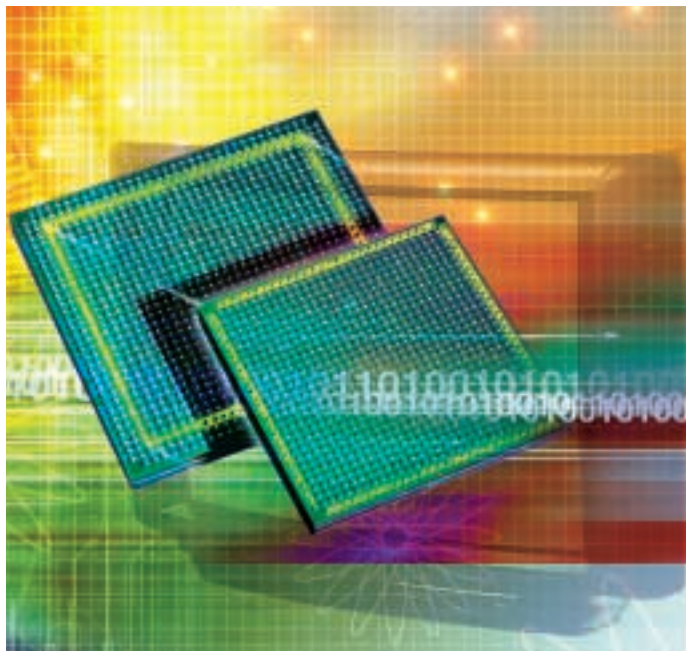
Decoding MPEG-2 Video

As devices and applications grow in complexity, the Amphion CS6651 MPEG-2 video decoder IP core can be implemented by taking advantage of the capabilities of Xilinx Platform FPGAs.

As part of a demonstration system, the core is implemented in a Xilinx Virtex XCV800 device. Including extra interfacing glue logic, and PAL/NTSC video encoder driver logic, the core consumes fewer than 8,000 slices and 26 block RAMs. The implementation benefited greatly from the following features of the target device:

- Ample high-performance block RAM

- Fast I/Os

- Extensive logic resources.

The functional blocks and simplified interfaces of the Amphion CS6651 MPEG-2 video decoder IP core are shown in Figure 1. The core requires a minimum clock speed of 27 MHz to maintain MP@ML decoding rates. Video elementary streams are accepted into the core via the byte-wide ES_Data input port on the elementary stream interface.

Parser

The front end of the core is the video elementary stream parser. It searches the syntax of the incoming stream for start codes at which decoding may commence. The parser extracts the various encoding parameters from the headers, which are used to direct subsequent decoding. The remaining variable-length encoded picture data is passed onto the VLC decoder.

VLC Decoder

Here, the Huffman-style VLC picture data is decoded. The outputs of this block include DCT block run-level codes and motion vectors for motion compensation of each macro block.

Run-Level Decoding and Inverse Quantization

The run-level decoder converts run-level codes from the VLC decoder into complete blocks of 64 quantized DCT coefficients. These coefficients are then converted from zigzag scan order to natural row order before being dequantized. Virtex block RAMs are used to support the scan conversion operation and the storage of custom quantization matrices, which may be sent in the elementary stream headers.

Inverse DCT

This unit performs the computationally intensive inverse DCT (IDCT) on the dequantized 8x8 blocks of DCT coefficients. Making use of the high-speed on-chip block RAM, the IDCT unit is capable of streaming data continuously, transforming in 64 clock cycles an 8x8 block of DCT coefficients into an 8x8 block of luminance or chrominance samples or prediction errors.

Motion Compensation and Picture Reconstruction

For each macro block in a P-picture or B-picture, the motion compensation unit takes the decoded motion vectors from the VLC decoder and translates them into row and column coordinates for the prediction samples in the reference picture. The frame store memory then requests these samples and retrieves them via the SDRAM interface. The samples are combined with other prediction samples, if necessary, to complete the prediction for the macro block.

The final stage of decoding is to add the prediction samples to the prediction error corrections from the inverse DCT unit and write the reconstructed samples into the frame store memory. If a block has no predicted samples, then the samples from the inverse DCT are the final samples and are passed straight through. Both motion compensation and picture reconstruction employ Virtex block RAMs for buffering predictions, samples, or prediction errors and final reconstructed samples.

SDRAM Interface

The frame store memory can be implemented using a commodity PC100 or better 64-Mb SDRAM part. The SDRAM interface handles the mapping of row and column motion compensation prediction requests and reconstructed sample writes into linear memory addresses. Motion compensation places particular memory bandwidth demands on the decoder implementation. To achieve adequate decoding performance, this unit must then arbitrate between the other functions, such as the display DMA, which also access the frame store.
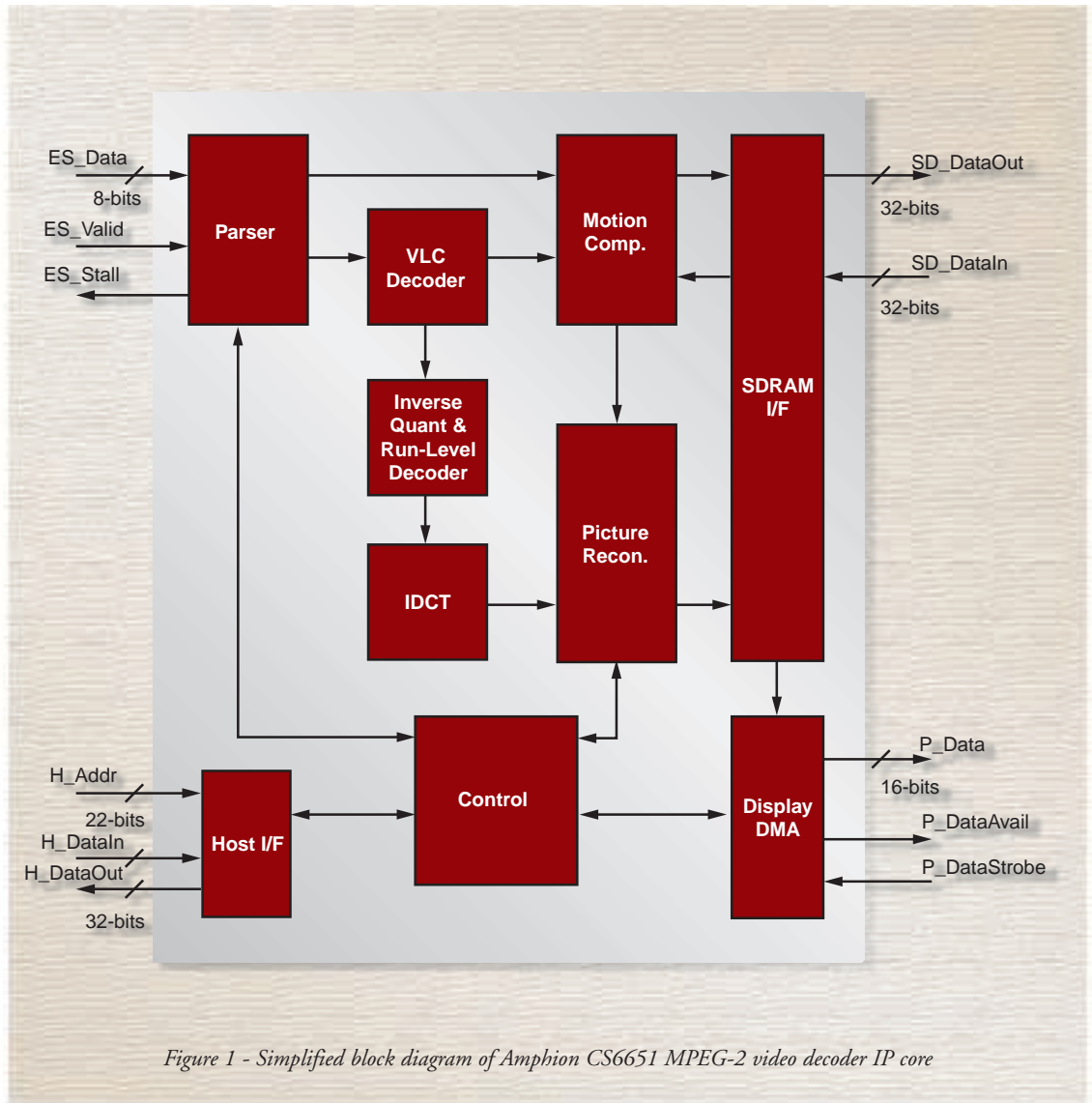

Figure 1 - Simplified block diagram of Amphion CS6651 MPEG-2 video decoder IP core

Display DMA

The display DMA unit retrieves decoded samples from the frame-store memory line-by-line for display. This unit has a configurable double-byte output interface for luminance and chrominance samples and can also perform chrominance upsampling to a 4:2:2 format in the vertical direction. This interface provides a number of handshake signals and flags (not shown in Figure 1) to easily allow for the addition of extra logic, to create sync pulses suitable for connecting the decoder to a NTSC or PAL video encoder chip.

Host Interface and Control

Access to internal control, status and video stream parameter registers within the core is provided via the host interface.

Simple 32-bit read/write access to the frame store is also available. In the demonstration system, a host processor controls the core to perform special effects modes such as pause and fast-forward in response to user commands.

Conclusion

MPEG-2 video decoding is likely to be an important feature of many future products. The Amphion CS6651 MPEG-2 video decoder IP core implemented on a single Xilinx Virtex Platform FPGA delivers MPEG-2 video decode capability with integrated SDRAM memory control and display DMA.

For more information about Amphion and its CS6651 video decoder, visit *www.amphion.com/cs6651.html*. ∑