TP 15.6 A Single-Chip MPEG-2 Video Audio and System Encoder

Govind Kizhepat, Kenneth Choy, Ronald Hinchley, Phillip Lowe, Roger Yip

iCompression, Inc., Santa Clara, CA

MPEG-2 encoding of motion pictures requires intensive computation and high I/O bandwidth to compress and packetize the video and audio streams for storage or transmission. As such, MPEG-2 encoding has traditionally been implemented with separate chips for video compression, audio compression, and system encoding. Such multi-chip solutions increase the total cost of MPEG-2 encoding and have consequently limited its use to expensive professional applications.

This single-chip MPEG-2 encoder integrates the video, audio, and system-level functionality for low-cost main profile at main level (MP@ML) encoding applications. Industry-standard interfaces reduce the total chip count and board space required for implementing PC-based and consumer-electronics applications.

With 6.1M transistors, the device is implemented in a 0.35μm 4-metal-layer CMOS process. It is packaged in a 420-pin ball grid array. The core operates at 3.3V, while the I/Os are 5V tolerant for interfacing with traditional TTL circuits. The device dissipates 3.5W at the 90MHz nominal operating frequency while processing video and audio streams, simultaneously. Table 15.6.1 summarizes the device specifications.

Figure 15.6.1 shows the block diagram of the device. The architecture is divided into the video and audio front-ends, and the system back-end. The video front-end consists of the video interface, video DSP, and video encoding engine. The audio front-end contains the audio interface and audio DSP. The system back-end includes the packetizer and multiplexer. The video DSP is shared between the video front-end and the system back-end. The video interface accepts 16b YUV from an external video A/D. It supports both 4:2:2 and 4:2:0 formats in both 720x480 (NTSC) and 720x576 (PAL) resolutions. The video interface converts 4:2:2 to 4:2:0 format by sub-sampling the U and V components. Horizontal and vertical filtering use 7-tap filters with fixed coefficients. The video encoding engine includes a complete encode-decode-reconstruct path with DCT, Q, IQ, and IDCT accelerators. It computes the motion vectors for frame and field pictures in a 3-step process illustrated in Figure 15.6.2. First, it performs a ±64pel horizontal and ±32pel vertical search on decimated pictures. It then narrows down the search location using a full pel search. Finally, it refines the search result by doing a half-pel search. The motion vectors along with the quantized coefficients are then encoded by the VLC. To balance the compression ratio and video quality, the encoding engine gathers statistics which are then used by the video DSP for computing the corresponding quantization parameters which determine the bit-rate of the encoded stream.

To meet the I/O bandwidth requirement during motion estimation, the video encoding engine interfaces with 8MB of external memory bank via a 64b wide data bus. This memory bank stores the current picture and forward and backward reference frame or field pictures. To reduce memory access latency during motion estimation, on-chip caches store the most-recently-used macroblocks and intermediate values. The encoded video stream is stored in a second 2MB external memory bank that is shared with the audio front-end for storing the compressed audio stream. The two memory interfaces allow the device to reach a combined maximum transfer rate of 1,080MB/s.

The audio front-end receives audio input through its I2S interface. The programmable audio DSP compresses the audio stream in either MPEG or Digital Dolby (AC-3) formats. The system back-end packetizes the compressed audio and video stream from the second memory bank by inserting the packet headers. It then multiplexes and encodes the packets in Elementary Stream, Transport Stream, or Program Stream formats. The encoded output stream from the device can be transmitted through a serial port, 8b parallel, or 32b PCI interfaces for communication, consumer electronics or PC applications.

Figure 15.6.3 shows the pipeline diagram of the 5-stage audio DSP. It implements a stack-based RISC architecture to reduce overall instruction code size. The DSP has a 24b multiplier and executes vector-type instructions to accelerate butterfly operation for audio encoding. Instructions and data for the DSP are stored in the second memory bank. To reduce memory access latency, the DSP has a total of twelve cache pages of internal SRAM. The stack register consists of 96 32b lines of SRAM. DSP instructions, such as audio compression algorithms and video quantization algorithms, can be down-loaded through the external ROM, microcontroller or PCI interfaces. The video DSP is similar to the audio DSP.

Figure 15.6.4 illustrates the system block diagram for PC-based and consumer-electronics applications. The PCI interface is used for PC-based applications. The host, ROM, and serial or parallel interfaces are employed for consumer-electronics applications such as recordable DVD.

The device contains five clock domains for the core, system, video, audio, and PCI interface. The system clock is controlled by an on-chip PLL and a balanced four-level clock tree. The worst-case on-chip clock skew is 115ps.

The design is implemented in a modular fashion using Verilog and is verified against a bit-accurate C-language model. Each major block is designed as a technology-independent, reusable module to allow easy technology porting and integration for consumer electronics applications. The highly-pipelined nature of the video compression path enables such methodology. Testability features include a full-scan implementation and general-purpose pins that monitor internal machine states. A chip micrograph is shown in Figure 15.6.5.

Reference:

G. Kizhepat, "Vivace-izC: The First Single-Chip MPEG2 Video, Audio and System Encoder", Microprocessor Forum 1998.

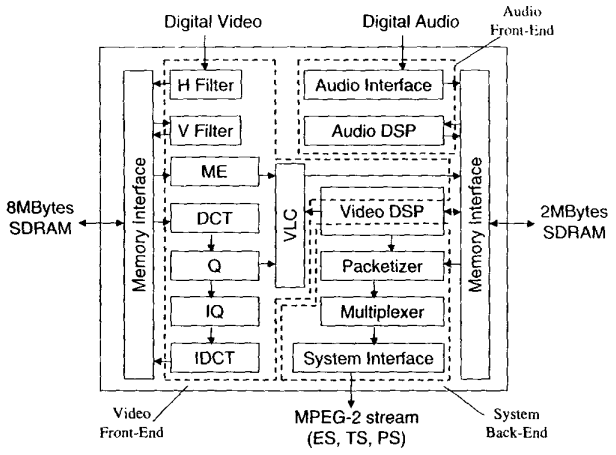

Figure 15.6.1: Block diagram.

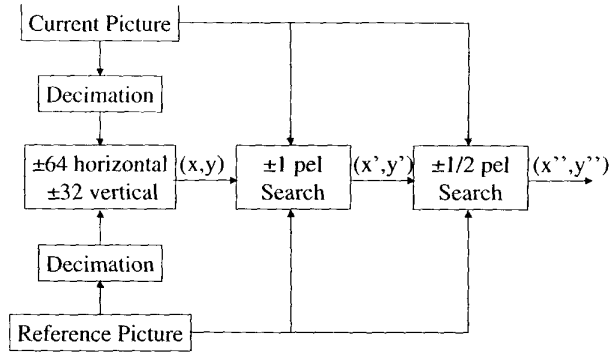

Figure 15.6.2: 3-step motion estimation.

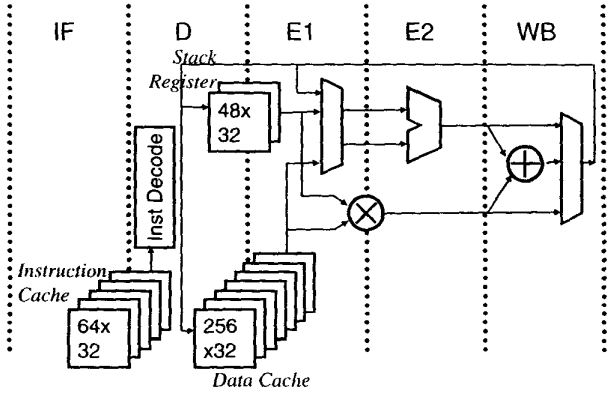

Figure 15.6.3: Pipeline diagram of audio DSP.

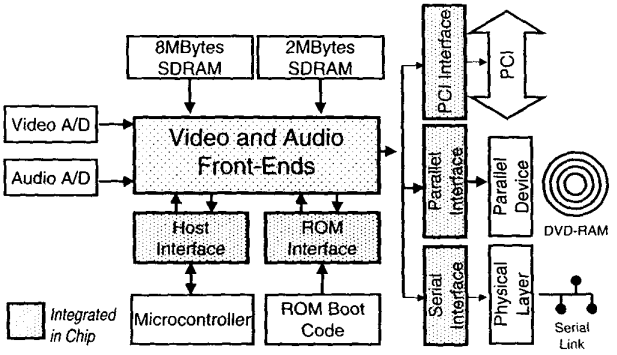

Figure 15.6.4: System block diagram.

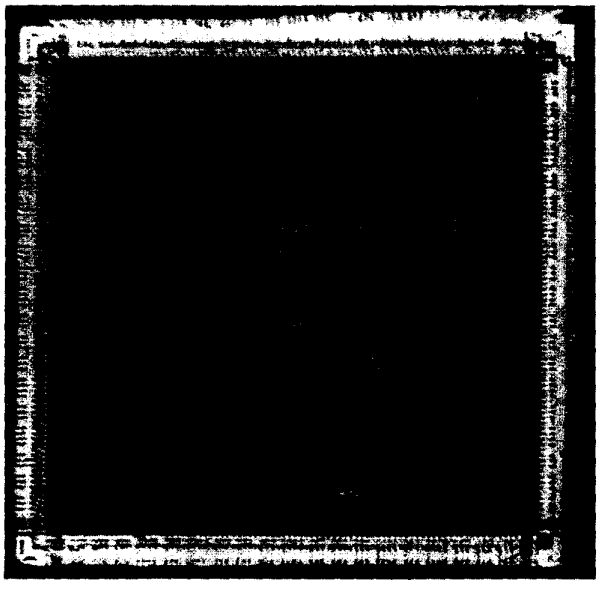

Figure 15.6.5: Chip micrograph.

Technology Metal 1 pitch Metal 2 pitch Metal 3 pitch Metal 4 pitch	0.35μm CMOS 1.12μm 1.32μm 1.32μm 2.00μm
Clock Frequency	Internal = 90MHZ Video = 27MHz Audio Input = 768kHz Encoded Output = 60 Hz PCI Interface = 33MHz
Power Supply	3.3V core: 5V-tolerant I/o
Transistor Count	6.1 million
Die Size	13.1mm x 13.1 mm
Package	420-pin Ball-Grid Array
Testability	Full-scan

Table 15.6.1: Device specifications