MICROPROCESSOR REPORT
THE INSIDERS' GUIDE TO MICROPROCESSOR HARDWARE
VOLUME 11, NUMBER 16
DECEMBER 8, 1997

DVx Sets New Standard for Digital Video

C-Cube Compresses Real-Time MPEG-2 Encoding Function Into Single Chip

by Peter N. Glaskowsky

Today, the digital-video market is divided into two camps. Applications that need low-cost video playback use decompression algorithms that can be implemented in inexpensive chips or host-based code. At the other extreme, video production and other applications that demand high quality and real-time encoding tend to favor more expensive solutions, such as Motion JPEG (MJPEG), that achieve only limited compression.

The Holy Grail of digital video has been to achieve television-broadcast quality at consumer price points. MPEG-2, especially at higher data rates, offers the necessary quality, but MPEG-2 encoding has always been prohibitively expensive for cost-sensitive markets like home video editing. Digital TV broadcasters like DirecTV use real-time MPEG-2 encoders that cost more than ten thousand dollars per channel, well beyond the reach of ordinary consumers and most video-content creators. With the introduction of the new DVx architecture by Les Kohn at October's Microprocessor Forum, C-Cube promises to remove this cost barrier, bringing real-time MPEG-2 encoding to the mainstream.

The DVx architecture will first appear in two products. The DVxpert 5110 is designed to support real-time encoding for MPEG-2's main level, main profile (ML@MP) and main level, professional profile (ML@PP) specifications, as used for DVD playback, digital-TV broadcasting, and video editing. Figure 1 shows the internal organization of the DVxpert 5110. The DVxpert 6210 is a two-chip set that provides extra processing power for professional video production work. Future DVxpert products will use even more DVx chips to support the higher resolutions and higher frame rates used by the advanced television (ATV) standards. ATV broadcasts are expected to debut in 1998, and broadcasters will need products like the DVx to produce the proper digital bitstreams. This is likely to be a good business for C-Cube and the other companies in this market, including Sony and IBM.

Meeting the Needs of the Digital-Video Market

A wide variety of digital-video algorithms is in common use today. The demands of Internet video distribution are very different from those of satellite TV broadcasting, while CD-ROM games and desktop video editing impose their own conflicting requirements. Different algorithms, or coder/decoder (codec) routines, have been developed to meet the needs of each market segment.

Because most digital-video production work is performed on personal computers and workstations, these systems are especially sensitive to the varying demands of the different market segments. Ideally, all video production should be performed using a single format that meets the needs of every target application, but to date no digital-video format has been able to accomplish this goal at a low cost.

Table 1 (see page 2) summarizes the current state of the art in digital video. Note that MPEG-2 meets the quality and playback-cost criteria while maintaining a reasonable data rate. This would make MPEG-2 the natural choice for video production work, except that MPEG-2 video encoding has always been either too slow or too expensive for affordable systems. Real-time (30 frames/s) encoding has required

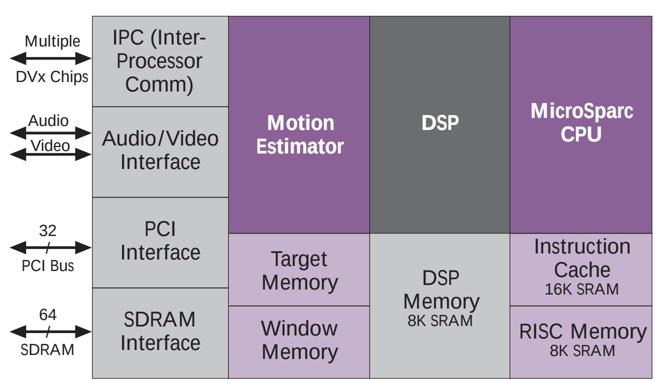

Figure 1. The DVx architecture includes a MicroSparc-based RISC CPU, a DSP, motion-estimation logic, and various other peripherals and peripheral interfaces.

Inside: Cyrix MXi ◇ StrongArm-1500 ◇ ARM9 ◇ Hyperstone ◇ Sparc64-III

expensive multichip solutions, while single-chip and software-based encoders take as long as a few seconds per frame, making them unsuitable for interactive video production.

Some media processors (see MPR 1/27/97, p. 10), notably the Philips TriMedia and Chromatic Mpact 2, are capable of MPEG-1 encoding. Mpact includes a motion-estimation engine, while TriMedia performs this function in software. Neither chip is today capable of real-time MPEG-2 encoding, but both companies are aware of the long-term value of this feature, and we believe both are likely to add it to future versions of their products.

Motion Estimation Is Strenuous

The key difference between MJPEG and MPEG-2 is that MJPEG encodes each frame into an independent unit that describes the complete frame, compressing individual 8×8 pixel blocks using the discrete cosine transfer (DCT) algorithm.

MPEG-2 encodes some frames this way (Intra or I-frames), typically one out of every 12, but other frames are defined by how they differ from these I-frames. Predicted, or P-frames, contain some new compressed blocks, but other blocks are defined by reference to the previous I- or P-frame. Bidirectional, or B-frames, consist entirely of references to the closest I- and P-frames before and after the B-frame (see MPR 6/17/92, p. 8).

C-Cube's Les Kohn describes the DVx, the world's first single-chip real-time MPEG-2 encoder.

The high cost of MPEG-2 encoders derives from the difficulty of computing these P-frame and B-frame references. Each reference describes how a 16×16-pixel or 8×16-pixel region can be derived from a similar region in another frame. To achieve the highest possible compression, MPEG-2 encoders must find the region that best matches the frame being encoded. To complicate the task, this match should be made on half-pixel boundaries. Each comparison requires calculating the sum of 256 or 128 absolute-difference calculations. An exhaustive search requires more than 170 trillion calculations per second for 720×480-pixel, 30-frame/s content, well beyond the capabilities of any current solution.

One final factor adds an additional level of complexity: the frames being searched must first be compressed, then decompressed, to produce the same bitmaps the MPEG-2 decoder will have available during playback. This means an MPEG-2 encoder must simultaneously perform most of the MPEG-2 decoding function.

	Low End	Production	Broadcast
Codec	Indeo, Cinepak	MJPEG	MPEG-2
Visual Quality	Low	High	High
Data Rate (Mbits/s)	0.5–1.5	20–50	4–8
Decoding Cost	Low	High	Low
Encoding Cost	Low	High	High
Encoding Speed	Slow	Real-time	Varies

Table 1. Current digital-video algorithms meet the needs of various markets, but none provides high quality, low data rates, and low-cost implementations.

Even with many parallel execution units, no existing MPEG-2 encoder (not even the DVx) can perform exhaustive motion-detection searches. Fortunately, this approach is not necessary or desirable. Instead of searching the entire image at half-pixel resolution, the encoder can achieve good results by searching at coarser increments for the closest match, then "zooming in" and searching around that area. This process, known as hierarchical motion estimation, can be repeated until a final match is found at half-pixel resolution. Also, it is not absolutely necessary to search the entire frame; few objects move more than a fraction of the total width or height of the frame between one frame and the next.

DVx Surpasses Multichip Products

Real-time MPEG-2 encoding hardware debuted in 1993 with the C-Cube 4600 family (see MPR 10/4/93 p. 20). Products based on these chips are used by DirecTV and other digital-TV broadcasters at a cost of roughly $12,000 per channel.

Prior to the debut of the DVx family, C-Cube's best MPEG-2 encoding product was the CLM4740, a seven-chip set rated at 35 billion operations per second (GOPS) for the motion-estimation task. This performance is sufficient to test the current block at offsets of ±100 pixels in the horizontal direction and ±60 pixels in the vertical direction for P-frames, or ±68 and ±36 respectively for B-frames.

The single-chip DVxpert 5110, with over 64 GOPS of throughput and a more efficient hierarchical motion-estimation algorithm, can search a ±202-pixel × ±124-pixel region for P-frames and ±100 × ±64 pixels for B-frames, permitting greater effective compression ratios and better image quality in the encoded video.

While the 5110 handles broadcast-quality video, professional users need even higher quality for production work, to ensure that repeated compression and decompression operations do not compromise the final result. Figure 2 shows the DVxpert 6210, a pair of DVx chips meant to handle this market. Even ATV formats such as high-definition television (HDTV) can be supported by using more DVx chips in parallel. C-Cube says real-time encoding of the 720×480-pixel, 60-frame/s baseline ATV standard will require two DVx chips, while the high-end 1,920×1,080-pixel specification (commonly known as HDTV) will require 8–10 chips for real-time compression.

The DVx chips have an 80-Mbyte/s interprocessor communication (IPC) bus that allows multiple devices to share the job of compressing high-resolution video streams. Each chip is assigned to a portion of the image, and the output from all of the chips is combined to create the final bitstream.

Because the DVx products are meant for use in interactive video-editing systems, C-Cube included the ability to decode two streams of MPEG-2 video simultaneously and process the two streams in the on-chip DSP to produce a single video output. This feature makes it possible to preview transition effects, such as dissolves and wipes, in real time. Also, unlike simple MPEG-2 playback applications, video-editing systems need the ability to directly access any random frame in the video stream. If the selected frame is a P- or B-frame, the necessary reference frames must be decoded first; the DVx architecture handles this requirement automatically.

The DVx is also capable of simultaneous MPEG-2 encode and decode operations, allowing it to decode a stream of source material, perform effects processing, and encode the result into a new MPEG-2 stream, all in real time.

Embedded MicroSparc Provides Flexibility

Unlike previous C-Cube MPEG encoders such as the Video-RISC Processor, or VRP (see MPR 1/27/97 p. 10), the DVx architecture uses a MicroSparc-based core, finally taking advantage of the SPARC license C-Cube acquired in 1995 (see MPR 11/13/95, p. 20). The core controls the other DVx processing elements through coprocessor interfaces. The core has 16K of instruction cache and 8K of SRAM for local data storage. Since the data set for the MPEG-2 operations performed by the SPARC core is a few hundred kilobytes in size, a data cache would not be very effective; instead, the local data memory must be explicitly managed by software.

The MicroSparc core provides far too little performance for the actual tasks of encoding or decoding MPEG-2 content. These tasks are handled by an on-chip DSP and motion-estimation logic. Both units are managed by the RISC core. The DSP unit handles filtering, DCT, quantization, and variable-length coding operations; it has 8K of dual-ported SRAM that can handle DMA transfers and simultaneous DSP operations.

The DSP unit is assisted by the motion-estimation coprocessor, which is capable of 64 absolute-difference operations per cycle at the chip's rated clock speed of 100 MHz. The motion-estimation algorithm is programmable, allowing the DVx to adapt its search pattern to the characteristics of the video stream being compressed.

The DVx can also adjust the amount of compression to achieve a specified target bit rate using one-pass or two-pass variable-bit-rate (VBR) encoding schemes. Some applications, such as DVD, impose separate limits on the instantaneous and average bit rates of the compressed video. For DVD, the instantaneous limit on video data is about 9 Mbits/s, while the average data rate must generally be low enough to fit the entire movie into the 4.7G capacity of a standard DVD disc. A two-hour movie must therefore have an average data rate of about 5.2 Mbits/s, including the audio tracks, subtitles, and other program content.

By using higher bit rates during complex scenes—crowds of people, choppy seas, explosions, and so on—and lower bit rates during easily encoded passages where the background is simple and the camera relatively stationary, the best overall quality is achieved within these limitations. The DVx's programmable RISC core can make some adjustments within a single compression pass based on scene complexity, or it can use the knowledge gained from one encoding pass to achieve roughly 25% better compression in a second pass.

Price & Availability

C-Cube has not announced pricing for the DVxpert chips but plans to ship silicon this month. The first DVx chips, the DVxpert 5110 and 6210, will be aimed at the professional video-production market, followed in 1998 by other DVx-based products for other markets. More information is available at www.c-cube.com.

High Integration Permits Low Price

C-Cube's Kohn says the first DVx chip, shown in Figure 3,

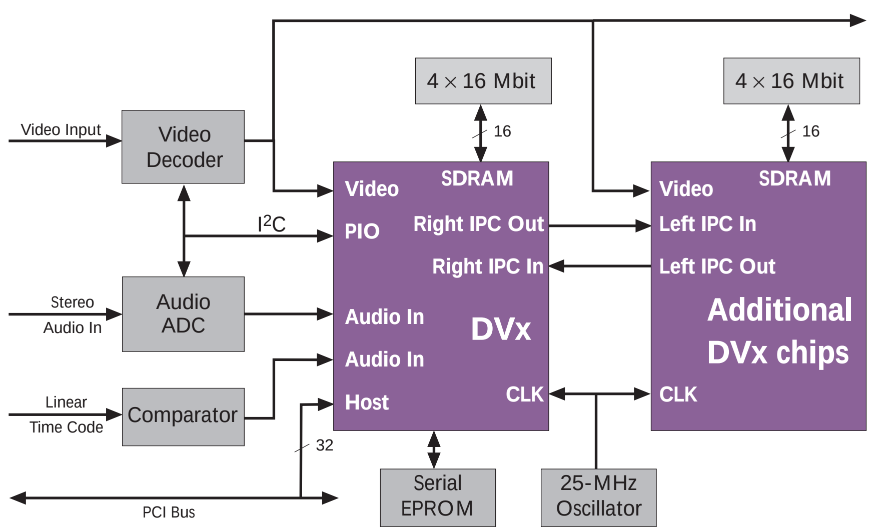

Figure 2. The DVxpert 6210 consists of two DVx chips operating in parallel to perform MPEG-2 encoding at higher-than-broadcast quality for video-editing applications. Each chip has local memory and shares processing tasks and peripherals through an interprocessor communication (IPC) interface.

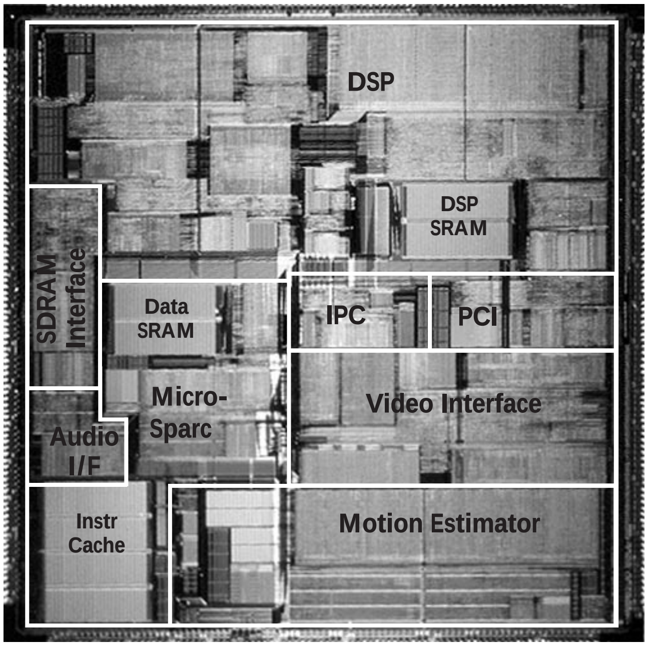

Figure 3. The DVx die, at 162 mm^2 in a 0.35-micron process, contains several different processing elements and interfaces.

will have a die size of 162 mm^2 in a 0.35-micron four-layer-and metal process. Packaged in a 352-contact BGA package and consuming 4.7 watts at 3.3 V, the DVx will be far less expensive than previous multichip MPEG-2 encoding products. According to the MDR Cost Model, a DVx chip costs about $70 to build.

While C-Cube has not announced pricing for the two initial DVxpert products, we expect the company to focus first on existing customers that have been paying thousands of dollars for equivalent performance. Once these high-margin applications have been satisfied, C-Cube can turn to the slightly higher volumes of the professional video-production market, and ultimately to the consumer space where sub-$1,000 video-editing boards will be used for corporate training films, home movies, and other more mainstream tasks.

We expect the company to adapt the DVx to a 0.25-micron process in 1998, reducing the die size to less than 100 mm^2 and the manufacturing cost to below $50. Boards based on such a chip could sell for $250 or less, reaching home computer users and even into the upper ranges of the consumer-electronics business. The new process should also increase the chip's clock speed to 150 MHz or so, reducing the number of DVx chips needed for professional editing or ATV systems.

In view of C-Cube's plan to drive the DVx into these cost-sensitive markets, it was necessary to provide a high level of integration on the new chips. Compared with previous VRP products, the new DVx chips add a PCI interface, audio and video ports, and the high-speed IPC interface. Complete PCI cards can be designed just by adding analog I/O for video and audio signals.

Consumer-electronics products based on the DVx are likely to need more on-chip peripherals, but not a PCI bus. We expect C-Cube to consider integrating audio and video A/D and D/A converters and related functions to help meet these needs.

C-Cube has already been fairly successful in the consumer market, with a large share of the Video CD market and several design wins for its MPEG Video Peripheral (MVP) chip in real-time MPEG-1 encoders from Dazzle, FutureTel, and others. MVP-based products sell for as little as $299 to PC owners who want an inexpensive way to capture digital video. Using MPEG-1 rather than MJPEG results in incoming data rates of around 1.5 Mbits/s, allowing these devices to be connected to the printer port and eliminating the need to install an expansion card.

Video CD players are widely used in Asia as an alternative to video-tape players. Many major manufacturers—including Panasonic, Samsung, Acer, and SMC—use C-Cube MPEG-1 decoders in their Video CD products. Video CD is popular in these markets because Video CD players and discs are less expensive and more reliable than VHS decks and tapes, advantages that DVD will eventually share.

DVx Meets Current Needs, Opens New Markets

The steady improvement of process technology is bringing us closer to the day when chips like the C-Cube DVx will be found in millions of homes, providing high-quality real-time compression of video for digital VCRs and future recordable DVD systems. We expect this to happen within just three to five years, depending on the rate of adoption of these new product types. Recordable DVD, in particular, is likely to become a huge market both in Asia, where it will provide a significant upgrade from today's Video CD products, and in the rest of the world, where it will ultimately displace both VHS and laserdiscs.

Given the ultimate size and importance of these markets, we expect C-Cube to face substantial competition. Sony and IBM already offer real-time MPEG-2 encoding products, but neither is as powerful or flexible as the DVx family. Sony is placing more emphasis on the digital video camera (DVC) standard, which uses an MJPEG-like algorithm to compress broadcast-quality video to a data rate of about 28 Mbits/s. DVC cameras from Sony and others have become popular products in the "prosumer" segment of the camcorder market, and DVC-based editing cards are starting to become available for PCs, but DVC is likely to be edged out by MPEG-2 products in a few years.

The arrival of real-time MPEG-2 compression will mark the convergence of the consumer and professional digital-video markets and lead to general support for broadcast-quality video creation, not just playback, on even the most affordable PCs. The DVx positions C-Cube to be a key player in these emerging markets. M