

WP 4.4: A Single-Chip MPEG2 Video Decoder LSI

Tatsuhiko Demura, Takeshi Ota, Kazukuni Kitagaki, Shun-ichi Ishiwata, Goichi Otomo, Shuji Michinaka, Seigo Suzuki, Nobuyuki Goto, Masataka Matsui, Hiroyuki Hara, Tetsu Nagamatsu, Katsuhiro Seta, Takayoshi Shimazawa, Kenji Maeguchi, Toshinori Odaka, Yoshiharu Uetani, Tadahiro Oku, Tomoo Yamakage, Takayasu Sakurai

Toshiba Corporation Kawasaki, Japan

This MPEG2 video decoder LSI decodes MPEG2 standard bit streams. The compression algorithm in the MPEG2 is based on discrete cosine transform (DCT), variable length coding, and motion compensation similar to the MPEG1, the earlier standard. However, the processing speed should be more than four times faster than MPEG1 [1, 2]. Moreover, several algorithms and structures to handle interlaced pictures are added to the MPEG1 standard. This LSI decodes in real time all motion-compensation modes and picture structures in MPEG2 bit streams of not only CCIR601 but also HDTV resolution, with distribution quality as shown in Table 1.

Operation is explained using the block diagram of the current chip in Figure 1. Compressed data supplied to the coded video data interface (CVIF) or the host interface (HIF) is written to a rate buffer (video buffering verifier; vbv buffer) in the external DRAMs through the FIFO and the 64b memory bus. Compressed data is re-read out from the rate buffer, and transferred to the VLD. Then, it is decoded in the VLD under control of the on-chip RISC. DCT coefficients are once stored in a buffer memory located in the run-length zigzag decoder (RZD) and other parameters are stored in a 16bx128w register file in the on-chip RISC. The on-chip RISC calculates actual motion vectors and addresses for the reference picture from extracted parameters, and sends them to the motion compensation unit (MC). The DCT coefficients are first converted by the RZD and then dequantized and transformed to real spatial domain data by the inverse quantization unit (IQ) and the inverse discrete cosine transform unit (IDCT), and stored in the data memory (DMEM). On the other hand, the reference picture addressed by the MC and the address generation unit (AGU) in the external memory are read out through the memory interface (MIF). If the motion vectors have accuracy of half pels, interpolation filtering is performed on the reference picture in the MC. If a macroblock is specified as a prediction type, an error component of image data stored in the DMEM and the reference data in the RMEM are added by the MC again, and the motion-compensated image data is obtained. Image data is rewritten to a display buffer located in external memory. Finally, the reconstructed image is re-read from the display buffer according to a specified interlaced or non-interlaced format and, if necessary, 4:2:0 to 4:2:2 conversion is executed in the display video data interface (DVIF) and read out.

There are three problems in MPEG2 decoder design. First, there is a memory bandwidth bottleneck from external DRAMs. Second, high-speed processing is needed due to the computation-hungry MPEG2 decoding. Third, the number of external DRAMs should be low to reduce system cost. The architecture improvements described below solve these problems.

The first is parallel operation of VLD and the RISC. A hand-crafted VLD macro has an auto-decoding mode besides normal decoding modes. The VLD can decode a certain length bit streams without aid of the RISC. This increases chip perfor-

mance and achieves decoding of the HDTV with distribution quality bit streams. Figure 2 shows a timing chart of the decoding sequences by the on-chip RISC and the VLD. In predictive coded macroblocks, reconstructing of motion vectors by the RISC is carried out in parallel with reconstructing of DCT coefficients by the VLD. In intra-coded macroblocks, decoding of dc coefficients is in parallel with decoding of ac coefficients by the VLD. This enhances decoding speed by a factor of about two while maintaining flexible bitstream interpretation using the normal decoding modes of the VLD.

The second improvement is on the DRAM interface. The rate buffer, the reference picture buffer and display buffer are located in external DRAM in a mixed way. This reduces the number of DRAMs and the number of pins. A memory bandwidth problem occurs, however, in the sharing of a 64b memory bus between reading from and writing to these buffers. To solve this problem, the BU (bus arbitration unit), several FIFOs, and memories are used to avoid data conflict on the memory bus. Figure 3 shows a state diagram of the algorithm in the BU. The basic idea of this algorithm is a combination of priority assignment and polling. Five requests for the memory bus are classified into the three priority groups as shown in Figure 3 and a grant to use the memory bus is given corresponding to the priority group. If requests from plural members in the first priority group are issued, the grant by the BU is allocated in a polling way. This type of control is required for the MPEG2 decoding but not for the MPEG1 decoding where the memory bottleneck is less severe. A verification of this algorithm was carried out by software simulation.

The last improvement is a data structure representing YCbCr data. To handle interlaced pictures that did not exist in MPEG1, eight Y pixels or a pair of four Cb and four Cr pixels are configured as a single 64b word when stored in the DRAMs as shown in Figure 4. This decreases effective bandwidth by about half compared to previous format. In the word configuration used, since only data of one field are contained in one word, there is no waste of bandwidth. This configuration is also suited for a macroblock form in the 4:2:0 picture format as shown in Figure 5. 32 luminance and 16 chrominance words correspond to the data in one macroblock.

Embedded arrays are used for the device. Hand-crafted macros are used for memories, the IDCT unit, and the VLD unit for cost-effective implementation. Random logic uses a gate-array approach. This approach provides the easy fast customization needed for this consumer LSI. Clock-tree synthesis based on a modified binary tree achieves 450ps clock skew by taking into consideration the capacitive unbalance of the macros by adjusting the length of each binary clock segment.

Device level features are summarized in Table 1. The chip is fabricated with a 0.5 μ m CMOS triple-level Al technology and is 15.0x15.0 μ m². Figure 6 is a chip micrograph.

References

- [1] Brinthaup, D., et al, "A Video Decoder for H.261 Video Teleconferencing and MPEG Stored Interactive Video Applications," ISSCC Digest of Technical Papers, pp. 34-35, Feb., 1993.
- [2] Sutarja, S., et al, "A 50MHz Vision Processor," IEEE Custom Integrated Circuits Conference, 1991

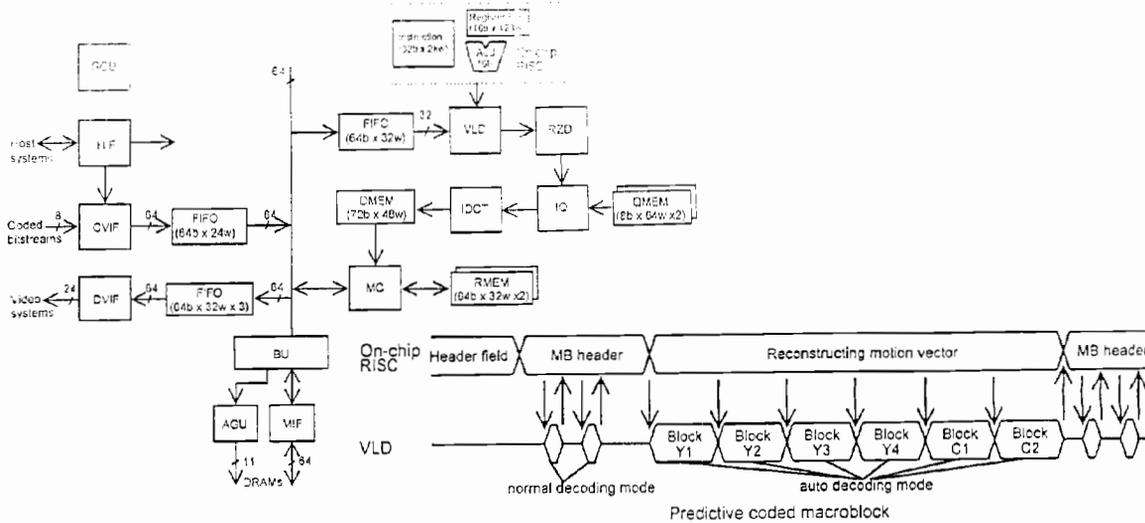


Figure 1: Block diagram.

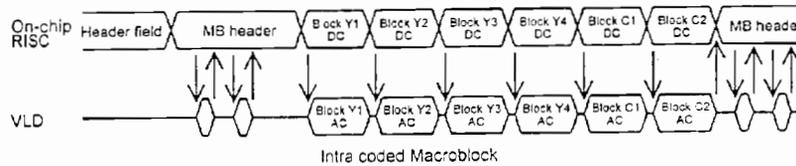


Figure 2: Decoding timing chart.

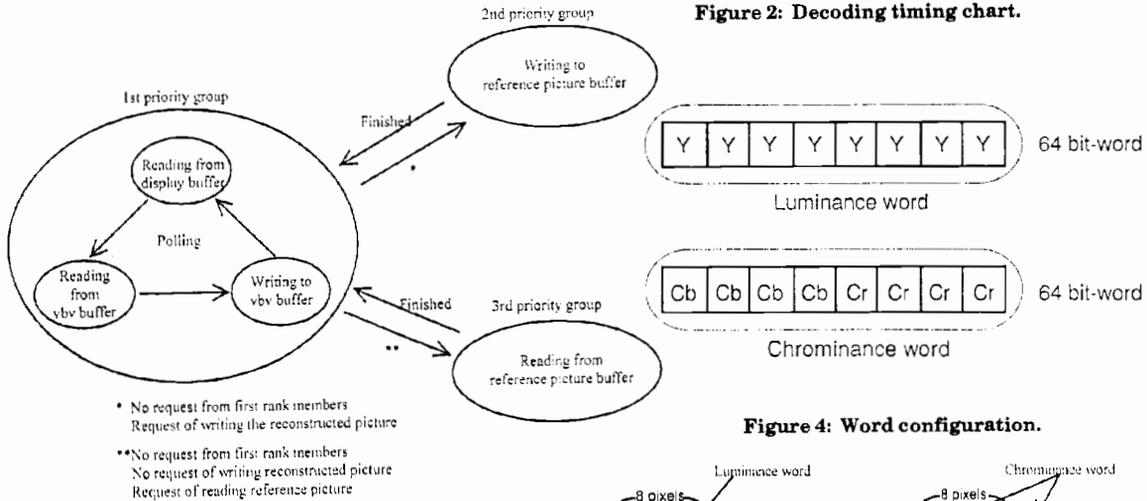


Figure 3: Bus arbitration state diagram.

Technology	0.5µm CMOS triple-level Al
Chip size	15.0x15.0mm ²
Number of transistors	1.1M
Random logic transistors	450k
Clock frequency	CCIR601(720x480) : 40MHz HDTV(1152x1024) : 70MHz
Power supply	3.3V
Package	299-pin PGA

Table 1: Device features.

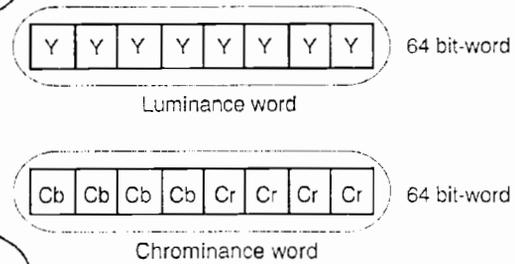


Figure 4: Word configuration.

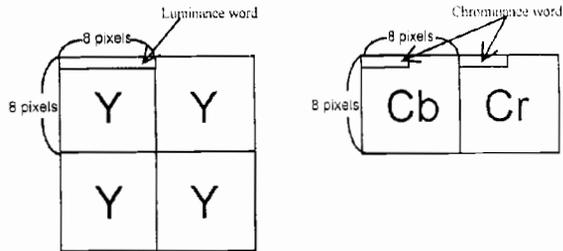


Figure 5: Macroblock.

Figure 6: See page 313.

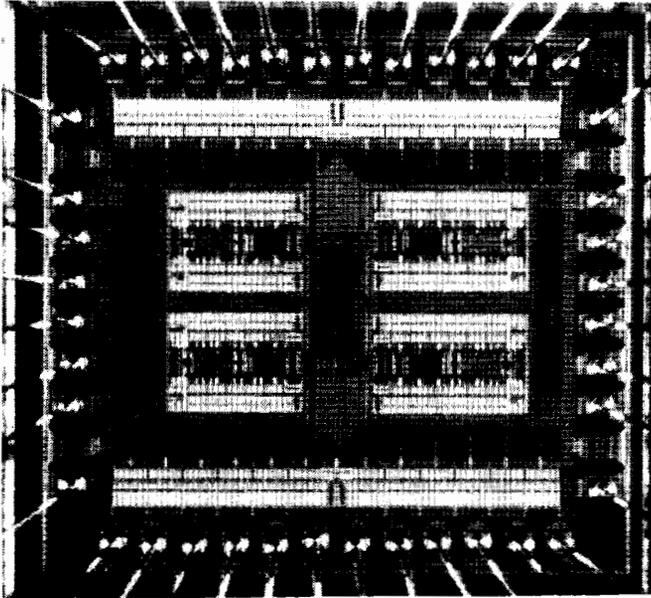


Figure 4: Chip micrograph.

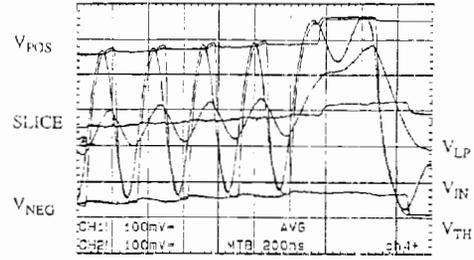


Figure 5: Experimental internal signals of slicer for 7Mb/s data pattern.

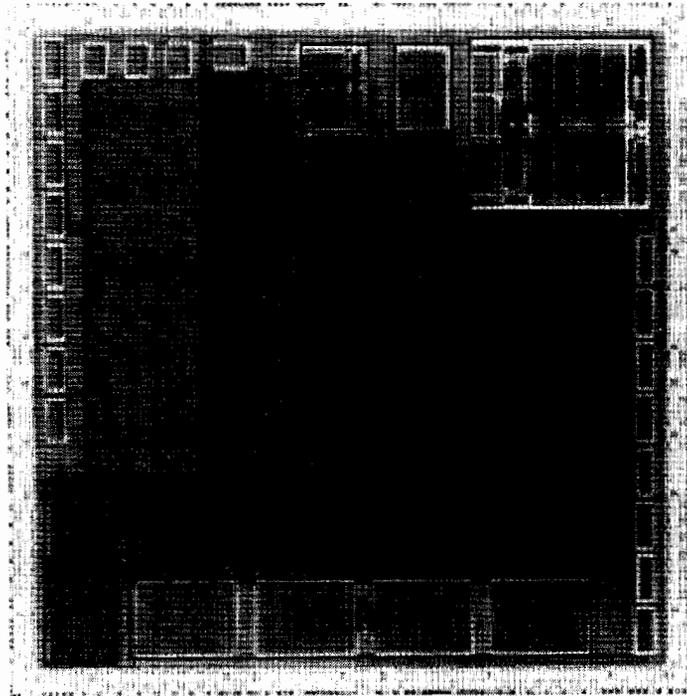


Figure 6: Chip micrograph.